













```
FT Intron 33976..34114
FT exon /*tag- h
FT exon 34115..34195
FT exon /*tag- i
FT Intron 34196..35759
FT exon /*tag- j
FT exon 35760..35901
FT Intron /*tag- k
FT Intron 35902..38781
FT exon /*tag- l
FT exon 38782..38996
FT exon /*tag- m

WO200058461-A1.
XX
XX PD 05-OCT-2000.
XX
XX PF 23-MAR-2000; 2000MO-EP02600.
XX
XX PR 26-MAR-1999; 99EP-0106343.
XX
XX PA (RAPP/) RAPPOLD-HOERBRAND G.
XX
XX PI Rappold-Hoerbrand G;
XX
XX DR WPI; 2000-656166/63.
XX
XX PT Novel nucleic acid sequence encoding human ataxia protein for screening
XX compounds useful for treating disorders relating to mutations in ataxia
XX gene
XX
XX PS Claim 6; Page 22-44; 47pp; English.
XX
XX CC The present sequence represents the human ataxia gene. The ataxia
XX protein and polynucleotides are useful for diagnosing and treating
XX disorders related to ataxia. Ataxia gene sequences are useful in
XX gene therapy, and as diagnostic tools or reagents for identifying and
XX characterizing genetic defect involved in the disorders and diseases
XX related to ataxia.
XX
XX SQ Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 other;

Query Match 13.9%; Score 227.8; DB 21; Length 39796;
Best Local Similarity 97.1%; Pred. No. 7.7e-34;
Matches 233; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 263 CCACATGACGACGCGGTGATGCTCTTGGCGCAACAGTGAACCCAGCGCTG 322
DB 33745 CCGCCAGGACGACGCGGTGATGCTCTTGGCGCAACAGTGAACCCAGCGCTG 33804
QY 323 CTTACGAGAAATCTGATGACTCTGTGACCTCGATCCCTCATGCTGACTATCT 382
DB 33805 CTTACGAGAAATCTGATGACTCTGTGACCTCGATCCCTCATGCTGACTATCT 33864
QY 383 GGAAGCGACGCTCTTGTGCTAATGAGAAAGGGCCAACTCCATGAGGTACCAAG 442
DB 33865 GGAAGCGACGCTCTTGTGCTAATGAGAAAGGGCCAACTCCATGAGGTACCAAG 33924
QY 443 ACAACAGTACTGCGCATCTTCAAGAAATGGAATGCTGTACAGATAGGCTGACC 501
DB 33925 ACAACAGTACTGCGCATCTTCAAGAAATGGAATGCTGTACAGATAGGCTGACC 33983

RESULT 6
ABI99254
ID ABI99254 standard; cDNA; 2404 BP.
XX
XX AC ABI99254;
XX
XX DT 07-MAR-2002 (first entry)
XX
XX DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:85.
XX
```

```
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KM vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
XX OS Mus musculus.
XX
XX PN WO200186188-A2.
XX
XX PD 22-NOV-2001.
XX
XX PF 18-MAY-2001; 2001MO-JP04192.
XX
XX PR 18-MAY-2000; 2000JP-0145977.
XX
XX (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
XX PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
XX DR WPI; 2002-034733/04.
XX
XX DR P-PDB; ABB57052.
XX
XX PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX expression levels of particular genes defined in the specification or
XX by determining the expression profile of a gene group comprising these
XX genes
XX
XX PS Claim 2; Page 265-268; 2690pp; English.
XX
XX CC The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (1) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (1). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (ABI99202 to ABI99912, encoding
XX the protein sequences in ABB57020 to ABB57374) or by determining the
XX expression levels of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischaemic condition-improving
XX drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
XX represent PCR primers for a mouse ischaemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX
XX SQ Sequence 2404 BP; 705 A; 493 C; 502 G; 704 T; 0 other;

Query Match 13.5%; Score 221; DB 24; Length 2404;
Best Local Similarity 56.5%; Pred. No. 1.5e-52;
Matches 503; Conservative 0; Mismatches 340; Indels 48; Gaps 3;

QY 170 GATATGATGCCAGGATTCGGCCCAATTTAAGGCCACCGCTGAACGTGACCTGCAACA 229
DB 279 GTTATGATGCCAGGATTCAGACCAACTTCAAGGCAATCTCTGTGATGATGATCAACA 338
QY 230 TCTTCATCAACAGTTTCAGTCCGTCACCAAGCACCAATGAGCTACCGGTGATGCT 289
DB 339 TTTTATTAATATTTTGGATTCATTCACAGCAACATGAGCTATAGGTTAAACATTT 398
QY 290 TCTTGGCGCAACAGTGAATGAGCCACGCT-GTCCACGAGAAATATCCTGA-----TG 343
DB 399 TCTTGAAGACAGAAATGAGATGACCCAGACTCAAGCTACTAGTACTTCAAGAGGCTGAG 458
QY 344 ACTCTGTGACCTCGATCCCTCCATGCTGAGACTGTATGGAAGCCAGACCTCTTTTG 403
DB 459 ATGCACTGACAGTTGACCCACCATGATATAGTGTGTAAGAACCTGACTATATCTTGG 518
QY 404 CTAATGAGAAAGGGGCAACTTCCATGAGGTGACCAAGCAACAGTACTAGCGCATCT 463
DB 519 CAATGAGAAAGTGCATTTTTCATGATGTGACCCAGAGAAATATCCTGTTTATATCT 578
QY 464 TCAGAGATGGGAATGTGCTGTACAGCATCAGGCTGACCTCATTTTCTCGCTGATGAG 523
DB 579 TTGCGGATGAGAGAGCTGCTGTGAGCATGAGCTGTATTAACATTTATGCTGCTG 638
QY 524 ACCTCAAGAACTTCCCATGAGACATCCAGACGTGACAGATGACGTTGAGAGCTCATCCA 583
```

```

Db 639 ACTTAACCTGTTCCCATGACACACAGCGTCAAAATGCAATTGAGAGCT----- 692
Oy 584 TACTGACAGCCCTCTGCATCTCTGTCACTTGGCTACACCATGAAGACCTCG 643
Db 693 -----TTGGATATATACACCGATGAATTTAA 716
Oy 644 TGTATAGTGGCTGGAAGATGCTGCTGCTCCAGAGGGCTGAGGGGCTGACTCTGCC 703
Db 717 GATTCACTGCGCAGTACAGAGATCTGTTCAAGTGAAGAAATGCTTACCTCAATTTG 776
Oy 704 AGTTATCTTGGCGGATGAGAGATCTAGGCTGTGTACCAAGCAGTACACAGAGGA 763
Db 777 ATATTAAAAAGAGATATCGAATATGCAACTGACAAATACTATAAGGCACGTGTT 836
Oy 764 AATTCACTGACATGAGTAAAGTTACCTGAGAGGAGGAGGCTACTATCATTC 823
Db 837 ACTACACTGTGTGGAGGTATCTTCACTGAGAGACAGGTGGGTCTCATATATGG 896
Oy 824 AGATGACATCCCAAGCTACTCATGCTATCCTGCTGGGTCTCTCTTGATCAACA 883
Db 897 GCGTATATGACCAACCTTGTCTGATGTTGTTCTCTCTGCTCTCTTCTTGATCAAC 956
Oy 884 TGGATCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 943
Db 957 CTGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1016
Oy 944 AGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
Db 1017 AGTGACACCACTGCGAGAGCGGAGCTTCTTAAATGCTTATGTAAGAGCGCTGATGT 1076
Oy 1004 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1054
Db 1077 GGCTCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127

RESULT 7
AAD22070 standard; DNA; 1614 BP.
XX
AC AAD22070:
XX
DT 12-FEB-2002 (first entry)
DE
DE Dermacentor variabilis clone 8 GABA-gated chloride channel DNA.
XX
KM Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;
KW GABA-gated chloride channel; recombinant expression; domestic animal; ds.
XX
OS Dermacentor variabilis.
XX
FH Key Location/Qualifiers
FT CDS 1..1614
FT /tag= a
FT /product= "GABA-gated chloride channel protein"
XX
PN MO200174884-A1.
XX
PD 11-OCT-2001.
XX
PE 28-MAR-2001; 2001MO-US09955.
XX
PR 31-MAR-2000; 2000US-193791P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Zheng Y, Cully D, Ludmerer S;
XX
DR WPI: 2002-010778/01.
XX
DR P-PSDB: AAE13312.
XX
PT New polypeptide useful for preventing or treating tick infestation, in
humans, dogs, cattle, horses, deer, or other wild or domesticated

```

```

PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid
PT (GABA)-gated chloride channel -
XX
PS Claim 8; Fig 3; 59pp; English.
XX
CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride
CC channels and their corresponding nucleic acid molecules. GABA-gated
CC chloride channel proteins and DNA's are useful for preventing and
CC treating tick infestation, particularly in humans, dogs, cattle, horses,
CC deer, or other wild or domesticated animals. The nucleic acids are useful
CC as hybridisation probes or Polymerase Chain Reaction primers for
CC identifying the presence of Dermacentor variabilis GABA-gated chloride
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic
CC acids are also useful for the recombinant expression of D. variabilis
CC GABA-gated chloride channel proteins. GABA-gated chloride channel
CC proteins exert toxic effects on other ticks or related parasites such as
CC mites. The present sequence is Dermacentor variabilis clone 8 GABA-gated
CC chloride channel DNA.
XX
SQ Sequence 1614 BP; 369 A; 475 C; 422 G; 348 T; 0 other;

```

```

Query Match 12.5%; Score 204.4; DB 24; Length 1614;
Best Local Similarity 55.1%; Pired. No. 7e-48;
Matches 426; Conservative 0; Mismatches 356; Indels 48; Gaps 3;

```

```

Oy 163 ACATCTGATATGATGCCAGATTTGCGCCCAATTTAAAGGCCCGGAGAGCTGACC 222
Db 154 ACAGTGGGTGAGACAGAGAGGTGAGGCCAAATTAATGCGGCTTCAAGTGAAGTTGGC 213
Oy 223 TGGACATCTTCAATCAAGATTTACGCTCCGTCACCAAGACCAATGAGTACCGGGTG 282
Db 214 GTCACTATATGAGATTTATACAGCATATGATACAGTCTGTAAGTACAAATGAGCTTACTTCT 273
Oy 283 AATGCTTCTTGGCGCAAGTGAATGACCCAGCGCTGTCTTACCGAGATATCTGTA- 341
Db 274 GACTTCTATTTCCGGCAATGCTGGCGGAGAGAGGAGGAGTCTGTTCCAGAAAGCCCGAGAC 333
Oy 342 --TGACCTCTGAGACCTGATCCCTGATCCCTGATCCCTGATCCCTGATCCCTGATCC 399
Db 334 CTGAGAGAGTGAAGTGTGGGCGCTGGAAGTGGCGGAGAGATCTGGGTACCGGACACTTTC 393
Oy 400 TTTGCTAATGAGAAAGGGGCAACTTTCATGAGTGAAGCAGGACCAACAAGTACTGCGC 459
Db 394 TTGCGCAAGAGAAAGAGCGGCTTCTTATGATGCGGCAACAAGCCCAACTTTCCTCCG 453
Oy 460 ATCTTCAAGAAATGGAGATGCTGTACAGATGAGCTGACCTCATTTTGTCTGCTG 519
Db 454 ATCGGCTCCGAGAGAGAGTTTTCGCCAGTATTCAGTGAAGTGAAGTGAAGTGAAGTGA 513
Oy 520 ATGACCTCAAGAACTTCCCATGAGACATCCAGACGTCGACGATGAGCTTGAAGCTCA 579
Db 514 ATGAGATCTGATATCTTCCGATGAGACAGCAAGCGTGCATCTATAGAGATGAAGACT-- 571
Oy 580 TCCATACTCTGACAGCCCTGTCCATCTCTGTCACTTTCAGTTGAGTGAAGTGAAGAC 639
Db 572 -----TTGGTTATATACATGAAGAAG 591
Oy 640 CTGCTGTTGATGCTGGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 659
Db 592 ATCCGCTACCGGTGGAGAGGTTTTCGCCAGTATTCGCTGCTGCTGCTGCTGCTGCTGCTG 651
Oy 700 CCCGAGTTTATCTTGGCGGATGAGAGATCTAGGCTGTTGTACCAAGCAGTCAACACA 759
Db 652 CCGGAGTTTCAAGTCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
Oy 760 GGGAAATTCACCTGATCGAGTAAAGTTTCACCTGGAAGGAGGAGATGGCTACTATCTG 819
Db 709 GGAAGTACTCCGCGCTGATGTAATGAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 768
Oy 820 ATTCAGATGATACATCCCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 879
Db 769 ATCCAGATGATACATCCCGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 828

```





PT DNA encoding alpha-4 and delta subunit(s) of the human GABA  
 PT receptor - also stably co-transfected eukaryotic cells expressing  
 PT receptors contg. these subunit(s), used for screening and designing  
 PT drugs  
 PS  
 PS Claim 10; Fig 3; 45pp; English.  
 CC A cDNA clone (AAT28989) codes for the delta subunit (AAR97299) of the  
 CC human GABA-A receptor. It was isolated from human hippocampus  
 CC cDNA using a probe generated from rat brain cDNA using primers (see  
 CC also AAT28996-97) based on the rat delta sequence. The cDNA can be  
 CC incorporated into a vector and used to direct expression of the  
 CC delta subunit in transfected cells, partic. rodent fibroblasts.  
 CC Co-transfected host cells capable of expressing the delta, alpha-4,  
 CC (see also AAR97298) and beta subunits can be used to screen or design  
 CC drugs that act on GABA-A receptors.  
 CC  
 XX Sequence 1555 BP; 289 A; 524 C; 466 G; 276 T; 0 other:  
 SO  
 Query Match 11.8%; Score 193; DB 17; Length 1555;  
 Best Local Similarity 53.3%; Pred. No. 1.3e-44;  
 Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;  
 Oy 142 CTGACAACTTATGGGCGCAACATCTGGATATGATGCGGAGATTCGGCCCAATTTTAA 201  
 Db 167 CTCCCAACCTGAGACGGCTGATAGCGGTACGCGCCCACTTCGGCTGGCATCGGA 226  
 Oy 202 GGGCCACCCGTAAGAGCTGCAACATCTTCATCAACATTCAGCTCCGACCAAG 261  
 Db 227 GGGCCCGGCTGATGATGCTGCTGCTGCTGAGAGGCGGACATGACACATCTCAAG 286  
 Oy 262 ACCAATGATGATCGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 321  
 Db 287 GCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 346  
 Oy 322 TCTTACCGAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 381  
 Db 347 TCTTACCGAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 403  
 Oy 382 TGGAGCCAGACCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 441  
 Db 404 TGGCTGCGGACGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 463  
 Oy 442 GACAACAGTATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 501  
 Db 464 GAGAACAGCTCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523  
 Oy 502 CTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561  
 Db 524 TCCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583  
 Oy 562 ATGACGCTGAGAGCTCATCTACTCTGACGCTGCTGCTGCTGCTGCTGCTGCTG 621  
 Db 584 CTGGACCTGGAGAGCT-----AC 601  
 Oy 622 GGCATACCATTAAGACCTCTGCTTGTAGTGGCTGGAAGATCTCTCTCTCAAGTG 681  
 Db 602 GCTTACTCATCGAGAGCATCTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTG 661  
 Oy 682 GCTGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 741  
 Db 662 CTGGACAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721  
 Oy 742 ACCAAGCATCTACAAACAGGAAATTCACCTGATGAGATTAAGTTTCACTGGAAG 801  
 Db 722 ATGAATTTCAAGTCGCTGGCCAGTTCCAGGCTGAGCTGACCTTCCAGCTGGGAG 781  
 Oy 802 CAGATGGCTACTATCTGATTCAGATGATGATGATGATGATGATGATGATGATG 861  
 Db 782 AACCGGGGGCTGATCATCTCAATCTTCAATCTTCTGCTGCTGCTGCTGCTGCT 841  
 Oy 862 TGGGTCTCTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATG 921

Db 842 TGGGTCTCTTCTGATCAACGACGAGGCGGCTGCGCCGACGAGGTGCTCTAGCATCAC 901  
 Oy 922 ACCGTGCTACCATGACACCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 981  
 Db 902 ACGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961  
 Oy 982 TAGTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1041  
 Db 962 GCGATCAAGGCACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1021  
 Oy 1042 CTGGAGTATGC 1052  
 Db 1022 GTGGAGTATGC 1032  
 RESULT 11  
 ABK27334  
 ID ABK27334 standard; DNA; 1359 BP.  
 AC  
 AC ABK27334;  
 XX  
 XX 09-APR-2002 (first entry)  
 DT  
 DT  
 XX  
 XX Mutant gamma-aminobutyric acid receptor gamma-2 subunit gene #3.  
 DE  
 DE Human; Anticonvulsant; Tranquilliser; Antianal; Antidepressant;  
 KW Nootropic; Neuroprotective; Neuroleptic; Antimigraine; Anorectic;  
 KW gamma-aminobutyric acid receptor subunit; GABA; epilepsy; anxiety;  
 KW manic depression; phobic obsessive symptom; Alzheimer's disease;  
 KW schizophrenia; migraine; obesity; receptor; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200198486-A1.  
 PN  
 PN 27-DEC-2001.  
 PD  
 PD 20-JUN-2001; 2001WO-AU00729.  
 PE  
 PE 20-JUN-2001; 2000AU-0008260.  
 XX  
 XX 20-JUN-2001; 2000AU-0008260.  
 PR  
 PR 13-SEP-2000; 2000AU-000098.  
 XX  
 XX 11-MAY-2001; 2001AU-0004953.  
 PA  
 PA (BION-) BIONOMICS LTD.  
 XX  
 XX Wallace RH, Mulley JC, Berkovic SF, Harkin LA, Dibbens LM;  
 PI  
 PI WPI; 2002-122280/16.  
 DR  
 DR P-PSDB; AU081289.  
 XX  
 XX Mutant gamma-aminobutyric acid receptor subunits and DNA molecule,  
 PT useful for diagnosing epilepsy, Alzheimer's disease, migraine, obesity,  
 PT anxiety, manic depression and schizophrenia -  
 PS  
 PS Claim 34; Page 73-75; 99pp; English.  
 PS  
 XX The invention relates to an isolated mammalian polypeptide (I), which is  
 CC a mutant of gamma aminobutyric acid (GABA) receptor subunit. The mutation  
 CC disrupts the functioning of an assembled GABA receptor, its functional  
 CC fragment or homologue, and creates a phenotype of epilepsy, anxiety,  
 CC manic depression, phobic obsessive symptoms, Alzheimer's disease,  
 CC schizophrenia, migraine and/or obesity. (I), the polynucleotide (II)  
 CC encoding (I) and antibody (III) to (I) are useful in the diagnosis of  
 CC epilepsy, anxiety, manic depression, phobic obsessive symptoms,  
 CC Alzheimer's disease, schizophrenia, migraine and/or obesity. (III) is  
 CC useful for treating the above conditions. (I)-(III) are useful in  
 CC screening of candidate pharmaceutical agents, where high-throughput  
 CC quantitative gene expression in biological samples. Oligonucleotides or  
 CC longer fragments derived from (II) are useful as probes in a microarray  
 CC used to monitor the expression level of large number of genes. (I)-(III)  
 CC are useful for the study of the function of a GABA receptor, to study the  
 CC mechanism of the disease as related to GABA receptor, for the creation of

CC explanted mammalian cultures which express a mutant GABA receptor and for  
CC the evaluation of potential therapeutic interventions. ABK27332-ABK27399  
CC represent mutant gamma-aminobutyric acid receptor subunit coding  
CC sequences and PCR primers of the invention.

SO Sequence 1359 BP: 254 A; 455 C; 405 G; 245 T; 0 other:

Query Match 11.6%; Score 189.8; DB 24; Length 1359;  
Best Local Similarity 53.1%; Pred. No. 9.8e-44;  
Matches 484; Conservative 0; Mismatches 382; Indels 45; Gaps 2;

```

Qy 142 CTAGACAATTATGGGGGCAACATCTGATATATGCGAGATTGCGCCCAATTAA 201
Db 121 CTCCCAACCTGGAGGGCTGATAGCGGCTACCGCCCAACTTCCGGCTGGCATGGA 180
Qy 202 GGGCCACCGGTGAGCGACCTGCAACATCTTCAACAGTTTACAGTCCGACCAAG 261
Db 181 GGGCCCCCGTGAATGAGCCCTTGGCCCTGGAGGTGGCCACATGACCCACATCTCAAG 240
Qy 262 ACCAATGAGACTACCGGTGAATGATCTTGGCGGCAACAGTGAATGACCGCATG 321
Db 241 GCCAATGAGATGATACATGATGAGTGTCTCTGACACAGACTGGCGGAGACAGGCTC 300
Qy 322 TCCATCCGGAATATCTGATGATCTTGGACCTCGATCCCTCATGCTGAGCTATC 381
Db 301 TCCATC---AACCAACCAACAGAGACCTGGGCTGAGACACCCCTTCTGTGACAACTG 357
Qy 382 TGAAGCCAGACCTCTTCTTGTATGAGAAAGGGGCAATCCATGAGGTGACCAAG 441
Db 358 TGGCTGCCCGACACTTCTATGATGATGACGCCAAGTGGCTGTCCACAGCTACAGGTG 417
Qy 442 GACAACAAGTTACTGCGCATCTTCAAGATGGGAATGCTGTACAGCATAGGCTGAC 501
Db 418 GAGAACAAGTCACTCCGGCTGACGCCGAGCGCGTGAATCTGTACAGATCCGAATCAC 477
Qy 502 CTCAATTTTGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 561
Db 478 TCCACTGTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 537
Qy 562 ATGAGCTGTGAGCTCATCTACTGATGATGATGATGATGATGATGATGATGATGAT 621
Db 538 CTGAGCTGTGAGACT-----AC 555
Qy 622 GGTACACACATGAAAGACCTGCTGTGATGATGATGATGATGATGATGATGATGATG 681
Db 556 GGTATCATGATGAGACATGCTGTACATGATGATGATGATGATGATGATGATGATGAT 615
Qy 682 GCTAGGAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 616 CTGGAACAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 675
Qy 742 ACCAAGCTACACACAGGGAATTCACCTGATGATGATGATGATGATGATGATGATGAT 801
Db 676 ATGAACCTTCAAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 735
Qy 802 CAGATGGCTACTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
Db 736 AACCGCGCTGATCATCTCATCTCATCTCATCTCATCTCATCTCATCTCATCTCATCT 795
Qy 862 TGGGTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 921
Db 796 TGGGTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 855
Qy 922 ACCGTCTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 981
Db 856 ACGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915
Qy 982 TACGTGAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1041
Db 916 GCCATCAAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 975
Qy 1042 CTGAGATATGC 1052

```

```

Db 976 GTGGAGTACGC 986
RESULT 12
AAD21378
ID AAD21378 standard; DNA; 1197 BP.
AC AAD21378;
XX 28-JAN-2002 (first entry)
DT
XX
XX R. sanguineus LGIC/GluC1 probe #1 to screen DVLGIC/GluC1 genes.
DE
XX Tick; acaricide; miticide; crop protection; nematocide; insecticide;
XX Lysand gated ion channel/L-glutamate gated chloride channel; LGIC/GluC1;
KW parasitic infection; pesticide drug screen; tick infection; probe; ss.
XX
XX Rhipicephalus sanguineus.
OS
XX W0200174899-A1.
XX
XX 11-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US09956.
XX
XX 31-MAR-2000; 2000US-193935P.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Cully DF, Zheng Y;
XX
XX WPI; 2001-662966/76.
XX
XX Novel purified nucleic acid molecule encoding Dermancentor variabilis
XX Lysand gated ion channel/L-glutamate gated chloride channel protein,
XX useful for identifying modulator compounds that are useful as
XX insecticides
XX
XX Example 1; Page 51-52; 98pp; English.
XX
XX The invention relates to a purified nucleic acid molecule encoding
XX Dermancentor variabilis (Dv) (American dog tick) lysand gated ion
XX channel/L-glutamate gated chloride channel (LGIC/GluC1) protein.
XX LGIC/GluC1 DNA is useful for identifying a compound that modulates
XX glutamate-gated channel protein activity and for identifying a modulator
XX of LGIC/GluC1 channel protein. LGIC/GluC1 DNA is used for establishing
XX novel insecticide screens, validate potential lead compounds with
XX insecticidal activity for use in treating parasitic infections in humans
XX or animals, as hybridisation probes to isolate related genes from other
XX organisms to establish additional pesticide drug screens and also used
XX in competition binding experiments or for functional chloride channel
XX assays to screen for compounds that activate, block or modulate the
XX channel. Heterologous expression of LGIC/GluC1 protein allows the
XX pharmacological analysis of compounds active against parasitic
XX invertebrate species relevant to animal and human health, especially in
XX the treatment of tick infestations directly related to Dv. LGIC/GluC1
XX protein acts as targets to identify modulators of the channels, where the
XX modulators act as effective insecticidal, acaricidal, miticidal and/or
XX nematocidal treatments for use in animal and human health and/or crop
XX protection. The modulators are useful for treatment or diagnosis of
XX specific disorders. The present sequence is Rhipicephalus sanguineus
XX LGIC/GluC1 1 probe which is used for screening Dermancentor LGIC/GluC1
XX genes.
XX
XX Sequence 1197 BP: 268 A; 358 C; 311 G; 260 T; 0 other:

```

Query Match 11.4%; Score 187.6; DB 22; Length 1197;  
Best Local Similarity 55.0%; Pred. No. 3.9e-43;  
Matches 466; Conservative 0; Mismatches 334; Indels 48; Gaps 3;

```

Qy 226 AACATCTTTCATCAGCTCCGCTACCAAGACCAATGAGTACCGGTGAAT 285
Db 97 AACATCTTGTAAAGAGTATGCGAAGATTGATGACGTCAACATGAGTACAGTCA 156

```





Db 475 TTTCTATG--GAAAGAGGGGATCTGTACAGGTACACAAAATCTCCACTTGGCCAGT 531  
 Qy 706 TTTATCTGGGAGATGAGAGGATCTAGGCTGTGTGTACCAAGCACTACACAGGGGAA 765  
 Db 532 TTCA---CGTGGAAAGGTTTCAACGACACTACGACAGTGGACCAAGCATGGCGAG 588  
 Qy 766 TTCACTGATCCAGTGAAGTTTCACTGAGAGGAGGAGTGGCTACTATCTGATTGAG 825  
 Db 589 TACAGCTGCTTGGCGGTGAGCTGGTGTTCAGAGCGAGTTTCAGCTACTACTGATTCAG 648  
 Qy 826 ATGTACATCCCACTGATCTACTGTCATCTCTGCTGGTCTCTTGTGATCAACATG 885  
 Db 649 ATTCATCATCCGCTGTGATCTGTCTGTCATCTGTCTGGTCTCTGCTGCTGACCCC 708  
 Qy 886 GATGCTGCTCCCTGGCCCTGTGGGCTGAGCATACACACCGTGTACCATCATGACCCAG 945  
 Db 709 ACCTGATCCCGGCGGCGAGTGTGCTGGGCTGACACACCTGTACCATGATGCCAGCAG 768  
 Qy 946 AGCTCTGGCTCCCGGCTCTTTGGCTAAGGTGTCTACGTGAAGCAATGACATCTGG 1005  
 Db 769 ATATCGGGCATACAGCGCTGCTGCTCCGCTTCTCTACCAAGCCATTTGAGCTGTGG 828  
 Qy 1006 ATGCTGTGTCTGCTCTTTGTGTGCTGCTGCTGCTGAGTATGCTCATTAATTTT 1065  
 Db 829 ACCGCGCTGTCTGTACCTGTGTATTCGGCGCTCTCTGAGTTGCGCTGTGATAC 888  
 Qy 1066 GTTCTCTG 1073  
 Db 889 GCCTCGCG 896

RESULT 14  
 AAD21395  
 ID AAD21395 standard; cDNA; 2138 BP.

AC AAD21395;  
 DT 28-JAN-2002 (first entry)  
 XX  
 DE R. sanguineus glutamate-gated chloride channel 1 cDNA clone, T12.  
 XX  
 KM Brown dog tick; glutamate-gated chloride channel; GluCl1;  
 KM GluCl2; crop protection; insecticide; nematocide; acaricide;  
 KM clone T12; ss.  
 OS Rhipicephalus sanguineus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 331..1683  
 FT /\*tag= a  
 FT /product= "R. sanguineus GluCl1 protein, T12"  
 FT /note= "This region is specifically referred in claim 7"  
 XX  
 PR WO200174838-A1.  
 PD 11-OCT-2001.  
 PF 28-MAR-2001; 2001WO-US09905.  
 PR 31-MAR-2000; 2000US-193934P.  
 PA (MERI) MERCK & CO INC.  
 XX Warmke JW, Yang Y, Cully DF, Hamelin MJ;  
 DR WPI: 2001-662863/76.  
 DR P-PSDB; AAE13037.  
 PT Novel L-glutamate-gated chloride channel proteins from Rhipicephalus  
 PT sanguineus for identifying compounds which modulate the channel  
 PT proteins, which are useful as insecticides, antihelmintics and  
 PT acaricides.

PS Claim 6; Fig 1; 89pp; English.  
 XX  
 CC The invention relates to Rhipicephalus sanguineus (brown dog tick)  
 CC L-glutamate-gated chloride channel proteins (GluCl1 and GluCl2)  
 CC and nucleic acid molecules encoding such proteins. GluCl1 channel  
 CC proteins are useful for identifying modulators. The compounds  
 CC identified as modulators are useful for insecticidal, mitacidal  
 CC and/or nematocidal treatment for use in animal and human health  
 CC and/or crop protection. The compounds are also useful in screening  
 CC for and selecting compounds active against parasitic invertebrate  
 CC species relevant to animal and human health, including worms,  
 CC fleas, ticks, mites and lice. Heterologous cell lines expressing  
 CC functional GluCl1 and GluCl2 channel functional forms are useful  
 CC for establishing functional or binding assays to identify novel  
 CC GluCl channel modulators. The present sequence is R. sanguineus  
 CC GluCl1 cDNA clone, T12.  
 SQ Sequence 2138 BP; 502 A; 661 C; 561 G; 414 T; 0 other.  
 Query Match 11.4%; Score 187.6; DB 22; Length 2138;  
 Best Local Similarity 55.0%; Pred. No. 5.4e-43;  
 Matches 466; Conservative 0; Mismatches 334; Indels 48; Gaps 3;  
 Qy 226 AACATCTTCATCAACAGTTTACGCTCCGTCAACAGACCAAGTACTACCGGTGAAT 285  
 Db 544 AACATCTTTGTAAGAAATATGCGGAGAAATGTATGACGTACACCATGAGTACAGTCA 603  
 Qy 286 GTCTTCTTGGGCAACAGTGAAGTACCCAGCGCTGTCTTACCGAGAAATATCTGATGAC 345  
 Db 604 ATGACGTTTCAGAGACAGTGGCGGAGCAGAGACATCCAGTACGACTTGGGCGGCGAG 663  
 Qy 346 TCTGTGACCTCGATCCCTCCATGCTGGAGCTCTATCTGAAAGCCAGACTCTTCTTGTCT 405  
 Db 664 GTTCTGCTACCTGACGCTTACCCGAGACGACAAAGCTTTGAAAGCCGAGACTCTTTTCTCC 723  
 Qy 406 AATGAGAAAGGGGCCCACTTCATGAGGTGACACAGGACAAAGTACTACCGCATTTTC 465  
 Db 724 AACGAGAAAGAGGACACTTCCACAAATCATCATGACCCCAAGTCTTCAAGCATACAT 783  
 Qy 466 AAGATGGGAATGTCTGTACAGCATCAGGCTGACCCCTCATTTTGTCTGCTGTATGAC 525  
 Db 784 CCCAAGGCGAGCTTCTTTCATGAGTATTCAGATTCCTTGGCTTTCATGTCCGATGAC 843  
 Qy 526 CTCAGACCTTCCCATGACATCCAGACATGACAGTACAGTGAAGCTCATCATCA 585  
 Db 844 CTGAATTTTATCTTGGATTAACAATCTCTCTATCTGATCATGTGATGCT----- 895  
 Qy 586 CTGTGACGCCCTTGTGCAATCTGTGTACCTTTGAGTTGCTACACCATGAAGACCTGTG 645  
 Db 896 -----ATGGGTATTAACAAGAGAGACCTGTG 921  
 Qy 646 TTTGAGTGGCTGGAAGATGCTCTCTCTCCCAAGTGGCTGAGGGGCTGACCTGTGCCCCAG 705  
 Db 922 TTTCTATG---GAAAGAGGGGATCTGTACAGGTCAACAAAATCTCCACTTGGCAGCT 978  
 Qy 706 TTTATCTGGGAGATGAGAGGATCTAGGCTGTGTGTACCAAGCACTACACAGGGGAA 765  
 Db 979 TTCA---CGTGGAAAGGTTTCAACGACACTACGACAGTGGACCAACATGCGCAG 1035  
 Qy 766 TTCACTGATCCAGTGAAGTTTCACTGAGAGGAGGAGTGGCTACTATCTGATTGAG 825  
 Db 1036 TACAGCTGCTTGGCGGTGAGCTGGTGTTCAGAGCGAGATTTCAGCTACTACTGATTCAG 1095  
 Qy 826 ATGTACATCCCACTGATCTACTGTCATCTCTGCTGGTCTCTTGTGATCAACATG 885  
 Db 1096 ATTCATCATCCGCTGTGATCTGTCTGTCATCTGTCTGGTCTCTGCTGACCC 1155  
 Qy 886 GATGCTGCTCCCTGGCCCTGTGGGCTGAGCATACACACCGTGTACCATCATGACCCAG 945  
 Db 1156 ACCTGATCCCGGCGGAGTGTGCTGGGCTGACACACCTGTACCATATGCCAGCAG 1215  
 Qy 946 AGCTCTGGCTCCCGGCTCTTTGGCTAAGGTGTCTACGTGAAGCAATGACATCTTG 1005



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2003, 19:53:28 ; Search time 4219 Seconds  
(Without alignments)  
11312.776 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgacaactctgtctcgc.....taaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBml:.\*  
1: gb\_ba:.\*  
2: gb\_htg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_sts:.\*  
28: em\_un:.\*  
29: em\_vl:.\*  
30: em\_htg\_hum:.\*  
31: em\_htg\_inv:.\*  
32: em\_htg\_other:.\*  
33: em\_htg\_mus:.\*  
34: em\_htg\_pln:.\*  
35: em\_htg\_rds:.\*  
36: em\_htg\_mam:.\*  
37: em\_htg\_vrt:.\*  
38: em\_sy:.\*  
39: em\_htgo\_hum:.\*  
40: em\_htgo\_mus:.\*  
41: em\_htgo\_other:.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1521.2	92.8	1874	6 AX392950	AX392950 Sequence
2	1503.8	91.7	1600	6 AX037565	AX037565 Sequence
3	1002.2	61.1	1251	10 AF462147	AF462147 Mus muscu
4	695.8	42.4	2045	5 DRE404970	AJ404970 Danilo rer
5	659.2	40.2	1384	10 RNGRSTR	Y00276 R.norvegic
6	656.2	40.0	1407	10 S73717	S73717 GLRA1-inhib
7	656	40.0	1715	9 HSGLYRA2	X52008 H.sapiens a
8	654.8	39.9	1359	10 RNO310837	AJ310837 Rattus no
9	654.8	39.9	3160	10 RNIGRA2	X61159 R.norvegicu
10	654.8	39.9	3865	10 RNNEOGLY	X57281 Rat NGlyR m
11	654.6	39.9	1350	10 RNO310834	AJ310834 Rattus no
12	654.6	39.9	2125	10 RATGLYRAL	D00833 Rattus norv
13	650.2	39.6	1857	9 HSGLYRAL	X52008 H.sapiens a
14	646.8	39.4	1744	4 AF268375	AF268375 Bos tauru
15	645.4	39.4	1431	10 S73718	S73718 Glra1-inhib
16	643.8	39.3	1374	10 RNO310835	AJ310835 Rattus no
17	643.8	39.3	1679	10 RNIGRAL	X55246 R.norvegicu
18	599	36.5	1222	10 RNO310836	AJ310836 Rattus no
19	585	35.7	2376	5 DRE308516	AJ308516 Danilo rer
20	575.6	35.1	1628	5 DRE5812	AJ005812 Danilo rer
21	572.6	34.9	1418	5 AY094974	AY094974 Morone am
22	552.8	33.7	3069	9 HSU93917	U93917 Human glyci
23	549	33.5	3085	5 AY094975	AY094975 Morone am
24	545	33.2	1443	10 RNO310838	AJ310838 Rattus no
25	543.4	33.1	2413	10 RATIGRAS	M55250 Rat inhibit
26	531.2	32.4	1682	5 AF488379	AF488379 Danilo rer
27	510.6	31.1	1392	10 AF362764	AF362764 Mus muscu
28	505.4	30.8	1595	9 DRE308517	AJ308517 Danilo rer
29	498.6	30.4	100269	5 HS1055C14	AL049610 Human DNA
30	366.2	22.3	657	6 E03608	E03608 DNA encodin
31	351.2	21.4	194822	2 AL671887	AL671887 Mus muscu
32	299	18.2	554	6 AX251654	AX251654 Sequence
33	244.8	14.9	2160	5 DRE404971	AJ404971 Danilo rer
34	239	14.6	1491	10 RNO310839	AJ310839 Rattus no
35	235.4	14.4	441	10 AF214575	AF214575 Mus muscu
36	232.6	14.2	1737	9 AF094754	AF094754 Homo sapi
37	232.6	14.2	1737	9 AF094755	AF094755 Homo sapi
38	232.6	14.2	2106	9 HSU33267	U33267 Human glyci
39	232.6	14.2	2649	9 BC032635	BC032635 Homo sapi
40	230.2	14.0	1661	4 AF268376	AF268376 Bos tauru
41	227.8	13.9	39796	6 AX037570	AX037570 Sequence
42	227.8	13.9	39796	9 HSU3563	Z93848 Human DNA S
43	223	13.6	940	9 AF018157	AF018157 Homo sapi
44	221	13.5	1465	10 MMGRBMRA	X81202 M.musculu
45	221	13.5	2404	6 AX305334	AX305334 Sequence

ALIGNMENTS

RESULT 1	AX392950	1874 bp	DNA	linear	PAT 23-MAR-2002
LOCUS	AX392950				
DEFINITION	Sequence 52 from Patent WO0212340.				
ACCESSION	AX392950				
VERSION	AX392950.1	GI:19700997			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	Yue, H., Thornton, M., Ramkumar, J., Tang, Y.T., Azinza, Y.,				
AUTHORS	Baughn, M.R., Yang, J., Yao, M.G., Lai, P., Walla, N.K., Ganahl, A.R.,				
	Hafalla, A.J., Nguyen, D.B., Patterson, C., Elliott, V.S.,				





Source	Organism	Reference Authors Title	Journal Title	Features	Source
Mus musculus.	Mus musculus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 1251)					
Groemer, T.-W., Becker, C.-M. and Becker, K.					
Localization of different glycine receptor isoforms in murine spinal cord					
unpublished					
2 (bases 1 to 1251)					
Groemer, T.-W., Becker, C.-M. and Becker, K.					
Direct Submission					
Submitted (21-Dec-2001) Biochemistry, University of Erlangen, Fahrstrasse 17, Erlangen 91054, Germany					
Location/Qualifiers					
1..1251					
/organism="Mus musculus"					
/strain="B6C3/Fe"					
/db_xref="taxon:10090"					
/chromosome="X"					
/tissue_type="spinal cord"					
1..1251					
/gene="Glr4"					
/note="mmGlr4"					
1..1251					
/gene="Glr4"					
/note="mmGlr4"					
/note="ligand-gated ion channel"					
/codon_start=1					
/product="glycine receptor alpha 4 subunit"					
/protein_id="AA169899.1"					
/db_xref="GI:18448711"					
/translation="MSPSDFLDKMGPSGYDARIRNFKGPPNVTNCFINSPGSGTETMDRVANVFLKQMDPRPLATREIPDSDLNPSSLESIMKPDLPFAEKANPHEVTDNDKLRLEKFNGLVLSIRLFLILSCPDNLNFPMDJOTMOLSEFYTNDMDLFEMLDAPAVOVAEGTLPOFILDERKDLGYCTKHVYTGKFTCLEVFLHLEKROMGYLILQMTIPSLIVLISWSPFIMMDAAPARVGLGITVLTMTQSSGRASLPKRYKRAIDIMAAVCLIEFPAALLEYAVNVFVSQHKFEMELRRRORRMEDIIRESREYFEGYGLCHLOARHGCGMESSISPPRPPLRLKGETMKRLYDRAKRIDTISRANVPPTFLVNITRYWVYKVLRSEDIHQAL"					
BASE COUNT	324 a	314 c	281 g	332 t	
ORIGIN					
Query Match	61.1%	Score 1002.2;	DB 10;	Length 1251;	
Best Local Similarity	87.2%	Pred. No. 8.9e-268;			
Matches 1128;	Conservative 0;	Mismatches 123;	Indels 42;	Gaps 1;	
124 ATGTCGCCCTTGATTTCTAGACAAATTTATGGGCGGACATGTGATATGATGCCAGG	183				
1 ATGTCGCCCTTGATTTCTAGACAAATTTATGGGCGGACATGTGATATGATGCCAGG	60				
184 ATGCGGCCCATTTTAAAGGCCCAACCGCTGAACGTGACCTCAACATCTTATCAACAGT	243				
61 ATTAAGGCCCAATTTCAAGGTCACCTGTGAATGTGACCTCAACATCTTATCAACAGT	120				
244 TTCAAGCTCCGACACAAGACACACATGACATACCGGGGGAATGTCTTTCGCGGAACAG	303				
121 TTGGGCTCCGTCATGAGACACCAATGACATACCGGGGGAATGTCTTTCGCGGAACAG	180				
304 TGAATGACCCAGCCGCTGTCTTACCGAGAAATATCTCTGATGATCTTGGACCTGCATCCC	363				
181 TGAATGACCCAGCGTGTGCGCTTACCGAGAAATATCAAGATGATCTCTTGGACCTAAACCA	240				
364 TCCATGCTGGACCTATATCTGGAAGCCAGACCTCTTCTTGTGAATGAGAAAGGGGCCAAC	423				
241 TCCATGCTGGACCTATATCTGGAAGCCAGACCTCTTCTTGTGAATGAGAAAGGGGCCAAC	300				
424 TTCCATGAGGTGACACGAGAAACAAGTTACTGGGCATCTTCAAGATGGGAATGTGCTG	483				
301 TTCCATGAGGTGACACGAGAAACAAGTTACTGGGCATCTTCAAGATGGGAATGTGCTC	360				
484 TACAGCATCAGGCTGACCTCATTTTGGTCCGTGATGGACCTCAAGAACTTCCCATG	543				
361 TACAGCATCAGATTCACCTCATTTTGGTCCGTGATGGACCTCAAGAACTTCCCATG	420				

OY		544	GACATCCGAGGTGCAGCATCAGCTTGAAGGGTCATCATCACTCGTAGGCCCTCTGCA	603
Dd		421	GACATACGAGACTGTACATCAGCTGGAAAGCT-----	454
OY		604	TCTCTGTACATTTCAGTTGAGGCTACAGCACTGAAAGACCTCGTTTGATGGCTGGAAGAT	663
Dd		455	-----TTGGCTTACACCATGAATGACCTCATGTTTGATGGCTAGGAAGAT	498
OY		664	GCTCTGCTGTTCGAAGTGGCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGCGGGATGAG	723
Dd		499	GCTCTGCTGTTCGAAGTGGCTGAGGGGGTGGACTCTTACCCCCAATTTATTTTGGGGATGAG	558
OY		724	AAGGATCTAAGCTGTGTTFACCAAGCACTACAACACAGGAAATTCACCTGCATGAGSTA	783
Dd		559	AAGGATCTAAGGTAATTGTACCAAGCACACTACAAATACAGSMAATTCACCTGATTGAGSTA	618
OY		784	AAGTTTCACCTGGAACGGGAGATGGGCTACTATCTGATTGATGATGATCATCCCAGCTTA	843
Dd		619	AAGTTTCACCTGGAAGCGGAGATGGGCTACTATCTGATTCAAAGTACATCCCAGCTTA	678
OY		844	CTCATATCGTATCCGTCCCTGGGTCCTCCCTTGAGAACCAATGAGATGCTGCCCTCGCCCGT	903
Dd		679	CTCATATCGTATCCGTCCCTGGGTCCTCCCTTGAGAACCAATGAGATGCGGCCCTCGCTCGT	738
OY		904	GTGGGCGCTGGGCACTACCAACCGCTCTCAACATGAGCAACCCAGACCTCTGGCTCCCGGGCC	963
Dd		739	GTAAGCGCTGGGCACTACCAACAGTCTCTCAACATGAGCAACTCGAGACTCTGGATTCAGGGCC	798
OY		964	TCTTTGCCCTAAGTGTCTCTACGTGAAGGCAATGCATCTGAGATGGCTGTGTCTCTC	1023
Dd		799	TCTTTGCCCTAAGTGTCTCTACGTGAAGGCAATGCATCTGAGATGGCTGTATGCTCTCTC	858
OY		1024	TTTGTGTCGCTGGCTTGCCTGAGATGATGCTGECCTAAATTTTGTTCCTCGCACCATAAA	1083
Dd		859	TTTGTATTTGCTGCTTGCCTTGTGAGTATGCTGCATCATTTTTGTCTCTCGCACCATPAG	918
OY		1084	GAATTCATATCAGCTTCGAAAGAAGCGACAGAGCGCCACAGCTTGAGAGAAATATGATCCAA	1143
Dd		919	GAATTCATATCAGCTTCGTAAGAAGCGACAGAGCTGACAGGTATGGAAGAAAGATCATTAAGA	978
OY		1144	GAAGTGTCTTTCTATTTCCGTGGCTATGAGCTTTGGGCCACTGCTCGACAGCAAGATGGA	1203
Dd		979	GAGAGCCGCTTCTATTTTTCGTGGCTATGAGCTTCAGACACTGCTTCAAGCAAGGGATGGA	1038
OY		1204	GCTCAATATGAAAGGTGTCGCAATTTATAGTCCCAACCTCCAGGCCCTCTTAAGGAA	1263
Dd		1039	GCTCAATATGAAAGGTTCAGCAATTTATAGTCCCAACCTCCAACCCCTCTTTCTAAAGAA	1098
OY		1264	GGAGAAACCAACGCGGAAACTCTACGTGGAGCTGAGACCAGAGAAATTGACACATCTCCGG	1323
Dd		1099	GGAGAAACCAATGAGGAAACTCTATGTGGAGCCGAGCAAGAAATTTGACACATCTCCAGG	1158
OY		1324	GCTGTCTTCCCTTTTCACCTTCTCATCTCTCAATATCTTCTACCTGGGTTGTCTATAAAGTG	1383
Dd		1159	GCTGTCTTCCCTTTTACTTTCTTGTCTTCATATCTTCTAATTTTGGTGTCTATAAAGTA	1218
OY		1384	CTATGTCAGAGAATATCCACAGGACTCTGGA	1416
Dd		1219	CTAAGGTAGAAGATATCCACAGGACTGTGA	1251
<b>RESULT 4</b>				
DRE404970				
LOCUS DRE404970 2045 bp mRNA linear VRT 21-NOV-2000				
DEFINITION Danio rerio mRNA for glycine receptor alpha2 subunit (glyr alpha				
gene).				
ACCESSION AJ404970				
VERSION AJ404970.1 GI:11322383				
KEYWORDS glycine receptor alpha2 subunit; glyr gene.				
SOURCE zebrafish.				
ORGANISM Danio rerio.				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				







S73717 1407 bp mRNA linear ROD 17-JUL-2000  
 LOCUS GLRA1-inhibitory glycine receptor alpha 1 subunit [mice, A/HeJ,  
 DEFINITION brainstem, spinal cord, mRNA, 1407 nt].  
 S73717  
 ACCESSION S73717.1 GI:765206  
 VERSION  
 KEYWORDS  
 SOURCE Mus sp.  
 ORGANISM Mus sp.; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1407)  
 AUTHORS Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L.,  
 Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P. and O'Connell, P.  
 A missense mutation in the gene encoding the alpha 1 subunit of the  
 inhibitory glycine receptor in the spasmodic mouse  
 Nat. Genet. 7 (2), 131-135 (1994)  
 JOURNAL 95004575  
 MEDLINE 7920629  
 PUBMED  
 REMARK Genbank staff at the National Library of Medicine created this  
 entry [NCBI g1bbseq 156492] from the original journal article.  
 This sequence comes from Fig. 2a.  
 FEATURES  
 source location/Qualifiers  
 1..1407  
 /organism="Mus sp."  
 /db\_xref="taxon:10095"  
 /tissue\_type="brainstem and spinal cord"  
 /note="A/HeJ"  
 1..1407  
 /gene="Glra1"  
 58..1407  
 /note="alternatively spliced; This sequence comes from  
 Fig. 2a; conceptual translation differs from that in  
 published reference; GLRA1 protein; Glyr alpha 1"  
 /codon\_start=1  
 /product="inhibitory glycine receptor alpha 1 subunit  
 short form"  
 /protein\_id="AAB32157.2"  
 /db\_xref="GI:9247215"  
 /translation="MYSFNTLRFYLMETIVFSLASKRPAARAPKPMSPDFLIDK  
 LMGRTSGDARLRPNFKGPVNVSCNITNFSGSIATETTMOTRVVIFLRQONDRLA  
 INEYVDDSLDPSMLDSIMKEDLFAENKGAHFEITDNLLEISRGNVLSIRI  
 TLTLACPMDLKNEPMDVQICIMLESFETNMDLFEQOQAVOVADGLTLPOLIK  
 EKDRLCYCTKHNTGKFTICIEARFHELEROMGYLLIOMYIPSLIYLISMSIWNDA  
 APARYGLITTYLTMTTOSGSRASLPKSYKALDIWMAVLLPFLVSLLEYAVNF  
 VSRHKEILRFRRKRHHKDEGEGRENFSAVGMPCIAOKOSISVGAANNNTTN  
 PPAKSKPEPRKRFIQRKKIDIKISRIQFPMFLIRMFYIITKIVREDVANK"  
 BASE COUNT 349 a 404 c 332 g 322 t  
 ORIGIN  
 Query Match 40.0%; Score 656.2; DB 10; Length 1407;  
 Best Local Similarity 72.6%; Pred. No. 2,2e-171;  
 Matches 942; Conservative 0; Mismatches 298; Indels 57; Gaps 5;

Db 395 TTGACCACTATATGTTGGATTCCATCTGGAGCCCTGACTTGTCTTTGCCAATGAGAAG 454  
 Oy 416 GGGCCAACTTCATGAGGTGACACGACCAACAACTTACGCGCATCTTCAAGATGGGA 475  
 Db 455 GGGCCACTTCCAGCAATCAGACGACGACACAACTGCTAGATCTCCCGATGGCA 514  
 Oy 476 ATGTCTGTAGACATCAGAGCTGACCCCTCATTTTCTCTCTGTGAGACCTCAAGAACT 535  
 Db 515 ATGTCTGTAGACATCAGAAATCAGACCTGACCTGCGCCCATGAGACCTGAAGAACT 574  
 Oy 536 TCCCATGACATCCAGACGTGACAGATGACACTTGAAGACTCATCTCTGACACC 595  
 Db 575 TCCCATGAGTGTACAGAGCTGTATCATGCAACTGGAAGCT----- 616  
 Oy 596 CTCTGCATCTGTCTACCTTTGAGTTGGCTACACCATGAAGACCTGCTTTGAGTGGC 655  
 Db 617 -----TTGATATACCATGACAGACCTCATCTTTGAGTGGC 652  
 Oy 656 TGAAGATCTCTCTGCTCCAAAGTGGCTGAGGGGCTGACTCTGCCAGTTTATCTTGC 715  
 Db 653 AGGACCAAG---GAGCTGTGAGGTGGAGAGGAGCTGACCTGCTGCTGATTTATTCGA 709  
 Oy 716 GGGATGAGAGATCTAGGCTTTGTACCAACACTACAAACAGGAAATTCACCTGCA 775  
 Db 710 AGGAAGAGAAAGTCTCAGATATCTCACCACCACTACAAACAGGTAATTCACCTGCA 769  
 Oy 776 TCGAGTAAAGTTTACCTGAGCAAGGAGATGGCTACTATGATGATGATGATGATGATG 835  
 Db 770 TCGAGGCCGATTTCCACTGCGAGGAGAGATGGCTACTACTGATCCAGATGTACATCC 829  
 Oy 836 CCAGCTTACTATCTATCTCTGCTCTGCTGCTCTCTGATCAACATGATGATGCTGCC 895  
 Db 830 CCAGCTGCTATCTATCTATCTCTGCTGATCTCTCTGATCAACATGATGATGCTGCC 889  
 Oy 896 CTGCGCGTGTGGGCTGGGATCACCACCGTCTCAACATGACCAACCCAGACTCTGGCT 955  
 Db 890 CAGCTGTGTGGGCTGGGATCACCACCGTCTCAACATGACCAACCCAGACTCTGGCT 949  
 Oy 956 CCCGGGCTCTTGGCTAAGTGTCTCTACCTGATGATGATGATGATGATGATGATGATG 1015  
 Db 950 CCCGAGCTCTCTTACCACCAAGTGTCTCTGATGATGATGATGATGATGATGATGATG 1009  
 Oy 1016 GTCTGCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075  
 Db 1010 GCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1069  
 Oy 1076 ACATTAAGAAATTCATGACACTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1135  
 Db 1070 AACCAAGA-----ACTGCTTCAATTTAGAGAGAGGAGGAGGAGGAGGAGGAGG 1123  
 Oy 1136 TCATCCAGAAAGTGTCTTCTATTTCCGCTGATGATGATGATGATGATGATGATGATG 1192  
 Db 1124 AGGTGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1183  
 Oy 1193 CAAGAGATGAGGATCCATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1249  
 Db 1184 CCAAGAGATGAGGATCTCTGCTGCAAGGATGAGGATGAGGATGAGGATGAGGATGAGG 1243  
 Oy 1250 CTCTCTTAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1309  
 Db 1244 GCGCATCCAGATCCCGGAGAGATGCGGAAGATCTCTCTCTCAAGAGAGGAGGAGGAG 1303  
 Oy 1310 ACACCATCTCCCGGAGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1369  
 Db 1304 ACAAGATATCTGAGATCGGTTTCCCATGAGGCTCTCTCTCTCTCTCTCTCTCTCTCT 1363  
 Oy 1370 TTGCTATTAAGTGTATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1406  
 Db 1364 TCATCTATTAAGATGCTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1400  
 RESULT 7  
 HSGLYRA2

LOCUS	HSLGTYA2	1715 bp	mRNA	linear	PRI 28-MAY-1993
DEFINITION	H.sapiens alpha-1 strychnine binding subunit of inhibitory glycine receptor mRNA.				
ACCESSION	X52009.1	GI:31850			
KEYWORDS	glycine receptor; inhibitory glycine receptor; strychnine binding.				
SOURCE	Homo sapiens.				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	Grønning-Joh, G., Schmieden, V., Schofield, P.R., Seeburg, P.H., Siddique, T., Mohandes, T.K., Becker, C.M. and Betz, H.				
AUTHORS	Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes				
TITLE	EMBO J. 9 (3), 771-776 (1990)				
JOURNAL MEDLINE PUBMED	90183975				
FEATURES	2155780				
source	Location/Qualifiers				
	1..1715				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="X"				
	/map="xp21.2-p22.1"				
	/tissue_type="brain"				
	/dev_stage="fetal"				
	297..1646				
gene	/gene="GLYRA2"				
	297..1646				
CDS	/gene="GLYRA2"				
	/note="strychnine binding alpha-1 subunit"				
	/codon_start=1				
	/product="inhibitory glycine receptor"				
	/protein_id="CA36258.1"				
	/db_xref="GI:31851"				
	/db_xref="SWISS-PROT:P23415"				
	/translation="MISFNTLRLLSGAVIFESLSASKKEAFNSATKPKMSPSFLDKLMGRPDSLDARIRNPFQGPVPVANSCTINFSEGSIAETMDVRNIPLROOMDPRLAYNEPDPDSLDIRNPFDMLSIIMKPDLEFANEKAHPHEITTDKLRLISINGWLVISIRITLTILACPMDLKNFMFDVOTCIQMLESFGYWNDLIFEQEGAVVADGLTPQFIILKEBDRYCTKHNYTKFKFCIARFHLEROMGYLYLOWIPSLILIVISLSEFWINDAPAVNGIGITVTVLMTMOSSGSRASILEKVSVKRVIDIMAKCLLFVSALILEVAANVPSROHKELRRRRKRHHKKDEAGEGRNFSAYGMCPRLACIDAKRGISYKGNNSNTNNPPRPSPSPPEMRKLFLORAKTIKISRIGFPMAFLIFNMFTWLITIKVREDVHNQ"				
BASE COUNT	424 a 480 c 407 g 404 t				
ORIGIN					
Query Match	40.0%; Score 656; DB 9; Length 1715;				
Best Local Similarity	71.9%; Pred. No.2-6e-171;				
Matches	952; Conservative 0; Mismatches 315; Indels 57; Gaps 5;				
OY	89	AGGAAGTCAAATCTGGAAACAAGGGGGTCCAGCCCATGTCCCCTTGATTTCCTAGACA	148		
Dd	367	AGGAGGCTGAAGCTGCCTCGTCCCAACCAAGCCTAATGTCAACCCTCGGATTTCCTGATA	426		
OY	149	AACCTATGGGGGGAACACTGTGGATATGATGCGAGGATCGGGCCCAATTTTAAGGCCAC	208		
Dd	427	AGCTTAATGGGGAACCTCCGGATATGATGCGAGGATCAGGCCCAATTTTAAGGTCGCC	486		
OY	209	CCGTGAACGTACCTGCAACACTCTTCATCAACAGTTTCAGCTCCGTCCACCAAGACACAA	268		
Dd	487	CAGTAGACGTGAGGTGCAACATTTTTCATCAACACAGCTTTGGTTCCATTCTGTGACACACA	546		
OY	269	TGGACTACCGGGTGAATGTCTTTTGGGGCAACAGTGAATGATGCCACCGCTGTCTAAC	328		
Dd	547	TGGACTAATAGGCTCAACATCTTCTCGCGGCGAGCAATGAGAAGACCCCCCGCTGACTATA	606		
OY	329	GAGATATACCGTGATGATCTCTGGACCTCGATTCCTCCATGCTGGAGCTATTTGGGAAGC	388		
Dd	607	ATGATATACCTGACGACTCTCTGGACCTGGACCCATTCATCTGTGAGCTCATTTGGGAAC	666		
OY	389	CAGACCTCTCTTGTCTAATGAGAAGGGGCAACATTCATGAGGTGACCGAGCAACA	448		

Db	667	CTGACCTGTTTTCCTCCACAGAGAGGGGGCCCTTCATGAGATACACCAAGACACACA	726
OY	449	AGTTACTGCGCATTTTCAAGAAATGGAGATGTGCTGTACACATCAGGCTCAGCTCATTTT	508
Db	727	AATTGTCTAAGGATCTCCCGGAATGGGAATGTCTCTACAGATCATGAAATCACCTGTACAC	786
OY	509	TGTCCTGCTGATATGGACCTCAAGAACTTCCCATATGACATTCAGACGCTGCACGATGCAGC	568
Db	787	TGGCGCTGCCCATGGACCTTGAAGAAATTTCCCATGATGTCCACACACTGTATCATGTCAAC	846
OY	569	TTTGAGACCTCATTCATCTCTGCGAGCCCTGTGCATCTCTGTACTCTTTCAGTTTGGCTTACA	628
Db	847	TGGAAAGCT-----TTTGATATTA	864
OY	629	CCATGAAGAGACCTTGCTGTTTGTAGTGGCTGGAGATGTCTCCTGCTGTCCAGTGGCTTGAGG	688
Db	865	CGAATGAATGACCTCATCTTTGAGTGGGAGGAGAACAG--GAGCCGTGCAGGTAGCAATG	921
OY	689	GGCTGACCTGCCCCAGTTTATCTTTGGCGGATGAGAGATCTAGGCTGTTTGTACAAAGC	748
Db	922	GACTTACTCTGCCCAAGTTATCTTGAAGAGAGAGAAAGAACTTGAGTACTGCACCAAGC	981
OY	749	ACTACAAACACAGGAAATTCACCTGCATCGAGTTAAAGTTTCACTTGGAACGGCAGATGG	808
Db	982	ACTACAAACACAGGAAATTCACCTGCATCTGAGGCCCGCGTTCCACTGAGACGGCAGATGG	1041
OY	809	GCTACTTCTGATTCAGATGTACATPCCCAAGCCCTACATCGATCATCTCTGCGGTCT	868
Db	1042	GTTACTACTCTGATTCAGATGTATATPCCCAAGCCCTGCTCATCTTGTATCTCTCATATGATCT	1101
OY	869	CCCTCTGATTCACATGATGATGCTGCCCCCTGAGCCCTGTGGGCTCTGGGATCACCAAGCTGC	928
Db	1102	CCCTCTGATTCACATGATGATGCTGCCCCCTGAGCCCTGTGGGCTCTGGGATCACCAAGCTGC	1161
OY	929	TCACCATGAGCACCCCAAGACTCTGTGGCTCCCGGGCTCTTTGCTTAAGGTCTCTAGCTCA	988
Db	1162	TCACCATGAGCACCCCAAGACTCTGTGGCTCTGTGGCTCTTTGTTGGCTCTCTGGAGT	1221
OY	989	AGGCATTCGACATCTGATGGCTGTGTGTGCTGCTTTGTTGGCTCTCTGGAGT	1048
Db	1222	AAGCATTTGACATTTGGATGGCACTTTGGCTGTCTGTGTCTGTCTGTGTCTGTGTGTGTGTGT	1281
OY	1049	ATGCTGCCATTAATTTTGTCTCTGTGATGACATTAAGAAATTCATPCACCTTCGAAGAAAGC	1108
Db	1282	ATGCTGCCATTAATTTTGT	1341
OY	1109	AGAGAGCCCAAGCTTGGAGAAATATTCATCCAGAAATGCTTTTATTTCCGTGGCT	1168
Db	1342	GGAACATTCACAAAGAGATGAACT-----GGAAGAGGCCCTTTAACTTCTCTGTCT	1395
OY	1169	ATGCGCTTGGGCT---ACTGCGCTGCAGGACAGATGAGTGGAGGTCCATGGAAGTTTC--TG	1222
Db	1396	ATGCGAGTGGGCGCCAGCCTGTCTTAAGAGCCCAAGAGATGTCATCTAGTCAAGGGGCGCAACA	1455
OY	1223	GCAATTTTATGCTCCCAACCTCCACCCCTCTTCTTAAGGAAGGAAGCAACCAAGCGGGAAC	1282
Db	1456	ACAGATTAACCAACCAACCCCTCTCTGACCATCTTAAGTCCCAAGAGAGATGGGAAC	1515
OY	1283	TCTTACGTGAGATGAGCCCAAGAAATTAACACATCTCCGGGCTGTCTCTCTTTCCTT	1342
Db	1516	TCTTCAATCCAGAGGCGCAAGAAATGACAAAAATATCCCGCATTTGGCTTCCCATGGCT	1575
OY	1343	TCTTCAATCTTCATATCTTCTACTGGGTTGTCTATAAGTGTATGTGCTCAAGAAATATTC	1402
Db	1576	TCTTCAATCTTCATATCTTCTACTGGGTTGTCTATAAGTGTATGTGCTCAAGAAATATTC	1635
OY	1403	ACCA	1406
Db	1636	ACAA	1639



```
RESULT 9
RINIGRAA2          3160 bp  mRNA  linear  ROD 23-MAR-1993
LOCUS              R.norvegicus mRNA for inhibitory glycine receptor alpha 2A subunit.
DEFINITION
ACCESSION          X61159.1 GI:288344
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus.
Rattus norvegicus.
Elkayota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 3160)
AUTHORS            Kuhse,J., Kuratov,A., Maulet,Y., Malosio,M.L., Schmieden,Y. and
                   Betz,H.
TITLE              Alternative splicing generates two isoforms of the alpha 2 subunit
                   of the inhibitory glycine receptor
JOURNAL            FEBS Lett. 283 (1), 73-77 (1991)
MEDLINE            91243883
PUBMED             1645300
FEATURES
source             location/Qualifiers
                   1..3160
                   /organism="Rattus norvegicus"
                   /db_xref="taxon:10116"
                   552..1910
                   /product="inhibitory glycine receptor alpha 2A subunit"
                   /protein_id="CAA43471.1"
                   /db_xref="GI:288345"
                   /db_xref="SWISS-PROT:P22721"
                   /translation="MNROLYNITLALFAFLGTHNHPREACCKDHSRSGKPSQTLSP
                   SDFLDLMGRISGTDARIRNENGPYVNTVNCNIFINSEGVETETMDYRNVLELRQW
                   NDSRLAISEYPDSDLDLPMDLSIMKPDLEFANEKANHDVETDKRLIRISNGRY
                   LYSIRLRLTSCPMDLKNEPMDVOTKMPDESGYTMNDLIFEMLSGPPVQVLEGLT
                   POFILKEKEKLGCTKHYNTGFTCLIEVKEHLEKROMGYKIDIMAVCLIFVPALE
                   WIMNDAPARVALGITVLTMTTOSGSRSLPKVSVYKALIDIMAVCLIFVPALE
                   YAAVNFYSROKHEFLRLRRORKRNEDEYRESRPSFGNGHCLQVYDGRAVKAT
                   PAMPLEPPPDADAIKKFFVDRAKRIDITISRAFPLEFLNFITWITYTIIRHEDVH
                   KK"
sig_peptide        552..632
mat_peptide        630..1907
BASE COUNT         902 a 669 c 645 g 944 t
ORIGIN
Query Match       39.9%; Score 654.8; DB 10; Length 3160;
Best Local Similarity 70.8%; Pred. No. 6.2e-171;
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;
QY 84 AAAAGAGGAAGTCAATCTGGACCAAGGGGTCCACCCCATGCTCTCTATTTCT 143
DB 635 AGACCAATGACTCCAGCTGTGAAAACATCCCTCGCAGACCCCTGCTCTCAGATTCTT 694
QY 144 AGACCAATGACTGTGGGAGCAATCTGATATGATGCCAGATTGGCCCAATTTAAAG 203
DB 695 GGAATAACTATGGGAGAGCAATCAGATATGATGCAAGAATCAGGCCAATTTTAAAG 754
QY 204 CCCAGCCGTGAGGTGACGCAACATCTTCATCAACAGTTTTCAGTCCGTCCACCAAG 263
DB 755 TCTCCAGTAAAGCTTACTTCGCAATATTTTATCAACAGTTTGGATGGGACAGAAAC 814
QY 264 CACAATGACTACCGGCTGATGCTCTCTTGGCGCAACAGTGAAGTACCACGCGCTGC 323
DB 815 CACCATGTGACTACCGATGAAACATTTTCTGAGACACAGTGGAAACGATTACAGCGCTGC 874
QY 324 CTACCGGATATCCGTATGACTCTGAGACCTGAGTCCCTCCATGCTGACTATCTG 383
DB 875 ATACAGTGAATACCCAGATATTCCTGGAATTTGGATCCATCGATGGAATTCGATTTG 934
QY 384 GAAGCCAGACCTCTCTTCTGATATGAGAAGGGCCCACTTCATGAGTGAACACGGA 443
DB 935 GAAACCGATTTGTTCTTTCGCAATGAGAAAGGCCAATTTTCATGATGTCACTGCA 994
```

```
QY 444 CAACAGTACTGCGCATCTTCAAGAAATGGGAATGTGCTGTACAGCATGAGCTGACCT 503
DB 995 TAACAGTGTGCTGGGATTTTCCAAAATGGCAAGTGCTCTACAGTATGTAGCTACCT 1054
QY 504 CATTTTCTGCTGCTATGAGACCTCAAGAACTTCCCATGAGATCCAGACGTGACGAT 563
DB 1055 GACTTATTCCTGTCGCCATGTGACCGCTGAAGAACTTCCAAATGATGTCCGACCTGTACAAT 1114
QY 564 GCACCTTGAGAGCCATCCATFACCTGTCCAGCCCTCCATCTCTGCACTTTCAGTTGG 623
DB 1115 GCAGCTGAGAGTT-----TTGG 1132
QY 624 CTACACCATGAAGACCTCGTGTGAGTGGCTGGAAGATGCTCTGCTGCCAAGTGC 683
DB 1133 GTACACCATGAATGACCTGATATTTGAGTGGTTAAGTATGATGTC---CAATACAGTGC 1189
QY 684 TGAAGGCTGACTGTGCCCATGTTATCTTGGCGGATGAGAAGATTTAGCTGTGTAC 743
DB 1190 TGAAGGACTCACCTGCTCAGTTATTTGAAAGAGAGAACCTTGCTATTTGCAC 1249
QY 744 CAAGCATACACACAGGAAATTCACCTGCATCGAGGTAAAGTTTCACCTGAGCGCA 803
DB 1250 AAGACATTAACACACTGSCAAGTTTACCTGCATTTAGAGTCAAGTTTACCTGAGCGCA 1309
QY 804 GATGGGCTACTATCTGATTCAGATGTACATCCCAAGCTCTACTGTCATCTGTCTG 863
DB 1310 GATGGGCTATTTATTTGACAGATGTATATCCAGGCTCTGATATTTGTTGCTG 1369
QY 864 GGTCCTCTCTGATTCACATGATGTGCTGCCCTCGCGGTGGCGCTGGGCAATCCAC 923
DB 1370 GGTCCTCTTTGGATTAACATGATGATGACGCTCCGCGCAGGGGTGCTTGATGCAACAC 1429
QY 924 CGTGCTACCATGACACACACAGAGCTGTGCTCCGGGCTCTTGGCTTAAGTGTCTTA 983
DB 1430 CGTGCTACCATGATGATACCCAGAGATGAGTTCCAGGCAATCTCTGCCAAGGTCTATA 1489
QY 984 CGTGAAGCAATGACATCTGATGCTGTGTCTCTCTTTGTGTCTGCTGCTTGTCT 1043
DB 1490 TGTGAAGCAATGACATCTGATGAGGAGTGTGCTCTCTTTGTGTCTGCTCTTACT 1549
QY 1044 GGATATGCTGCCATTAATTTGTGTTCTGTCAGCATTAAGAAATTCATAGACTGTGAAG 1103
DB 1550 GGAATATGACAGCACTGACTTGTCTCAGGCAACACAGAGATTCCTTGCTGCTCGGAG 1609
QY 1104 AAGCAGAGCGCGCCAGACGCTTGGAGAAATATCATTCACAAGAGTCTTATTTCCG 1163
DB 1610 ACAGCAAGAGAGGAGATTAAGAGAAAGATGTACTCGTGAAGAACCGCTTTAACTTCAG 1669
QY 1164 TGCTATGCTTGGCGCAGTGCCTGCGACGCAAGATGAGAGTCAATGAGAGTCTGG 1223
DB 1670 CGGTTATGGATGGTCTACTGCTCCCAAGTGAAGATGTTACAGCTGTCAAGGCTAC--- 1726
QY 1224 CATTTATAGTCCCAACCTCCAGCCCTCTTCAAGGAAGGAGAAACCGGGAACCT 1283
DB 1727 -----ACGTGCCAACCCACTTCCGCAACCCCAAGATGCAAGTCTTCAAGAAAGAA 1780
QY 1284 CTACGTGACTGAGCAAGAAATTAAGCAATCTCCCGGCTGTCTTCCCTTCACTTT 1343
DB 1781 GTTTGTGATCGGGCAAAAGAAATCGACACATATCTCGAGTGGCTTCCCATCGGCTT 1840
QY 1344 CCTCATCTTAATATCTTCTACTGCGTGTCTTAAGTCTTATGCTGACAGATATCCA 1403
DB 1841 CCTCATTTTCAACATCTTACTGATGATACATACAGATCATTCGGCATAGAGATGCCA 1900
QY 1404 CCAG 1407
DB 1901 CAAG 1904
RESULT 10
RINNEOGLY          3865 bp  mRNA  linear  ROD 07-AUG-1991
LOCUS              Rat Nglr mRNA for neonatal glycine receptor.
DEFINITION
```

ACCESSION X57281 GI:56743  
VERSION glycinine receptor; NG1YR gene.  
KEYWORDS Rattus norvegicus.  
SOURCE Rattus norvegicus.  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 3865)  
AUTHORS Hishinuma, F.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUN-1991) F. Hishinuma, Mitsubishi Kasei Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan  
REFERENCE 2 (bases 1 to 3865)  
AUTHORS Akagi, H., Hirai, K. and Hishinuma, F.  
TITLE Cloning of a glycine receptor subtype expressed in rat brain and spinal cord during a specific period of neuronal development  
JOURNAL FEBS Lett. 281 (1-2), 160-166 (1991)  
MEDLINE 91200276  
PUBMED 1707830  
FEATURES  
source Location/Qualifiers  
1..3865  
/organism="Rattus norvegicus"  
/strain="Wistar"  
/db\_xref="taxon:10116"  
/tissue\_type="spinal cord"  
/dev\_stage="10 days old"  
1..3865  
/gene="NG1YR"  
1..3865  
/gene="NG1YR"  
/evidence="experimental"  
1173..2531  
/gene="NG1YR"  
/codon\_start=1  
/product="neonatal type of glycine receptor"  
/protein\_id="CAA0549.1"  
/db\_xref="GI:56744"  
/db\_xref="SWISS-PROT:P22771"  
/translation="MNROLVILALFAFLGTNHFREARCKDHSRKHPSOTLSP SDFDLKMGRTSGDARIIRPNKGPVNVNINIFINFGSVTEETTMGYRNIERROOM NDSRLAISEYPPDSLDLDPSMDSIMKRPDLFEANERKANPHDVTDKLRLISNGRY LYSIRLRLTSCPMDFKPNPDVOTCTMOLESFYTNDLFEWLSGPVAVAGTLV PQLTKEKEKELGCTKHNYGKFTICVKEHLEKROMGYLLIOMYIPSLIVLISWVF WIMNDAPARVALGELITVLTMTTOSGSRASLPVSYVKADIMAAVCLLFVALLLE YAAVNFVSRHKEFLRLRRORONKEEDYRESRFNPSGGMGHCLQVGDGTAVKAT PANPLPQPPDADAIRKKKFDRAKRIDTISMAAPLALFLFNIFWTIKIRHEDVH KK"  
BASE COUNT 1087 a 833 c 889 g 1056 t  
ORIGIN  
Query Match 39.98; Score 654.8; DB 10; Length 3865;  
Best Local Similarity 70.88; Pred. No. 6.4e-171;  
Matches 938; Conservative 0; Mismatches 332; Indels 54; Gaps 3;  
OY 84 AAAAGAGAACTCAATCTGGAAACCAAGGGTCCAGCCCATGTCCCTCTGATTTCT 143  
DB 1256 AGACCATGACTCCGCTGTGAAACCAACCTCCGACACCTGTCTCTCAGATTCTT 1315  
OY 144 AGAACAATCTATGGGGGGAACATCTGATATGATGCCAGATTTGGGCCAATTTAAAG 203  
DB 1316 GGATTAACCTAATGGGAGGACATCAGATGATGAGCAAGATCAAGGCAATTTTAAAG 1375  
OY 204 CCCACCGGTGAACCTGACCTGACATCTTCATCAACAGTTTACGTCCTCGCACCAAG 263  
DB 1376 TCCCTCCAGTAACGCTTACTTCTCATATTTTATCAACAGTTTGGATCGGTCACAGAAC 1435  
OY 264 CACAATGACTACCGGGTGAATGTCTCTTGGCGCAACAGTGAATGACCCACCGCTGTC 323  
DB 1436 CACCATGTACGACGAGCAATTTTCTGAGACAGAGTGAAGCAATTCACGCGTGC 1495  
OY 324 CTACCGAATATATCTGATGACTCTCGACCTGATCCCTCCATGCTGGAGCTATCTG 383

DB 1496 ATACAGTAGTACCAGATGATTCCTGTGATTTGGATTCATCGATGCTGATTCGATTTG 1555  
OY 384 GAAGCCAGACCTCTTCTTCTGTAAGAAAGGGGCAACCTTCATGAGTGACACAGCA 443  
DB 1556 GAACCAAGATTTGTTCTTTTGGCAATGAGAAAGAGCCAAATTTTCATATGTACACACTGA 1615  
OY 444 CAACAACTTACTGGCATCTTCAAGAAATGGAAATGCTCTGTACAGATCAGGCTGACCT 503  
DB 1616 TAACAAAGTGTGGGATTTCCAAAATGCAAAAGTGCTGTACAGATTTAGACTCACCTT 1675  
OY 504 CATTTTCTGCTGCTGATGAGACCTCAAGAACTTCCCATGACATCCAGACGTGACAGAT 563  
DB 1676 GACTTTATCTGTCCATGAGCCGGAAGAACTTTCCAAATGAGATGTCACACCTGCAAT 1735  
OY 564 GCACCTGAGAGCTCATTCATCTCTGACAGCCCTGTCATCTGTACATTTGAGTTG 623  
DB 1736 GCACCTGAGAGCTT-----TTGG 1753  
OY 624 CTACACCATGAAAGACCTGTGTTGAGTGCTGGAAGATGCTCTGCTCCAACTGCG 683  
DB 1754 GTACACCATGATGATCCGTGATTTGAGTGTGTTAAGTATGATGTC---CAGTACAACTTGC 1810  
OY 684 TGAGGGGCTGACTTGTCCCACTTATCTTGGGATGAGAGATATAGGCTTTGTAC 743  
DB 1811 TGAAGGACTCACCTGCTCAGTTATTTGAAAGAGAGAGAACCTTGCTATTTCCAC 1870  
OY 744 CAACCACTACAAACAGAGAAATTCACCTGATGAGAGTAAAGTTTCCACCTGGAAGGCA 803  
DB 1871 AAGCACTTACAACTGAGCAAGTTTACCTGATGAGATCAAGTTTACCTGGAAGGCA 1930  
OY 804 GATGGGCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 863  
DB 1931 GATGGGCTATATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1990  
OY 864 GGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 923  
DB 1991 GGTCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2050  
OY 924 GGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 983  
DB 2051 GGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2110  
OY 984 GGTGAGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1043  
DB 2111 TGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2170  
OY 1044 GGAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103  
DB 2171 GGAATATGCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATG 2230  
OY 1104 AAGCAGAGGCGCCCAAGCTTGGAGAGATATCATCAAGAAAGTGTCTTCTTCTTCCG 1163  
DB 2231 AAGCAGAGAGGCGCAAGATTAAGAGAGATGATGATGATGATGATGATGATGATGATG 2290  
OY 1164 TGCTATGCTTGGGCACTGCTGCAAGCAAGATGAGATGAGATGAGATGAGATGAG 1223  
DB 2291 CGTTATGAGGATGAGTCACTGCTCAAGTGAAGATGATGATGATGATGATGATGATG 2347  
OY 1224 CATTTATAGTCCCAACCTCAGCCCTCTTCTTAAGGAGAGAGAAACACGCGGAAT 1283  
DB 2348 -----ACTGTCACACCTTCTCGCAACCCCAAGAGATGCAATCTTATCAAGAGAA 2401  
OY 1284 CTACGTGAGTGAACCAAGAAATTTGACACCATCTCCGCGGCTCTTCTTCTTCTTCT 1343  
DB 2402 GTTTGTGATGGGCAAAAGAAATGACACCATATCTTCGAGCTGCTTCCACATGCGCTT 2461  
OY 1344 CTTATCTTCAATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1403  
DB 2462 CTTATCTTCAATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1461  
OY 1404 CCAAG 1407  
DB 2522 CAAAG 2525

RESULT 11	LOCUS	RNO310834	1350 bp	mRNA	linear	ROD 15-AUG-2002
DEFINITION	Rattus norvegicus mRNA for glycine receptor alpha 1 precursor,					
ACCESSION	AJ310834					
VERSION	AJ310834.1 GI:13548654					
KEYWORDS	alternative splicing; glycine receptor alpha 1 precursor.					
SOURCE	Norway rat.					
ORGANISM	Rattus norvegicus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;					
TITLE	Rattus.					
JOURNAL	1					
MEMLINE	Barto, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.					
REFERENCE	Openings of the rat recombinant alpha 1 homomeric glycine receptor					
AUTHORS	as a function of the number of agonist molecules bound					
TITLE	J. Gen. Physiol. 119 (5), 443-466 (2002)					
JOURNAL	2 (bases 1 to 1350)					
COMMENT	Groot-Kormelink, P.J.					
FEATURES	Direct Submission					
source	Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of					
	Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,					
	London, WC1N 1AX, UNITED KINGDOM					
	related splice variants AJ310835 and AJ310836.					
	Location/Qualifiers					
	1..1350					
	/organism="Rattus norvegicus"					
	/strain="Sprague-Dawley males"					
	/db_xref="taxon:10116"					
	/tissue_type="Whole Brain"					
	/clone_lib="rat Brain marathon-ready cDNA"					
	/dev_stage="8-12 weeks"					
	1..1350					
	/function="ion channel1"					
	/note="alternative splicing, primary transcript"					
	/codon_start=1					
	/evidence="experimental					
	/product="glycine receptor alpha 1 precursor"					
	/protein_id="CAC35978.1"					
	/db_xref="GI:13548655"					
	/translation="MYSFNTLRFYIMTIVEFSLAASKEDAAARSAPKPSPEFLD					
	LMGRSGIDARIRPKPPVNVSGNITINFSGIAETTMRYRNIPIRQOMNPRLL					
	INERPDSDLDPSMLDSIMKPDLEFAEKCAHFELITDMLKLSKNGVLSIR					
	TLTLACPMDLKNFMPDVOTCIQMOLESEFYNNDLFEWQEGAVQADGLTPQIL					
	BEKDIRYCTKHNKNTGKFCIEARFLEKQMCYLYLIOMYPSLILVILSIFWIMJMD					
	APARGLGITVLTWTFTSSGSRAFLPKVSYVKAIIDWAVCLLFEFALLEEYAVNV					
	VSRQKELLFRFRKRHRHKDEGGEGGRNFSGYGMGPACLDKDISVKGANNNTT					
	PAPASRSPKPEMRKIFLORAKKIDKISIGFPMALFLINMFTWIIIKYVRDEDYHK					
BASE COUNT	332 a	395 c	318 g	305 t		
ORIGIN						
Query Match	39.9%	Score 654.6;	DB 10;	Length 1350;		
Best Local Similarity	72.6%	Pred. No. 6-2e-171;				
Matches 941;	Conservative 0;	Mismatches 299;	Indels 57;	Gaps 5		
OY	116	CCGAGCCCATGTC	CCCCCTCTGATTTCT	PAGACAACTATGGGCGGAACATCTGATATG	175	
DB	98	CCAACTATGTC	CACCTCTCGACTTCT	CGAATTAATGGAAGGACACTTCTGGGTATG	157	
OY	176	ATGCGAGGTTGGGGCC	CAATTTTAAGGGCCG	CCGCGTGAAGTACGACATCTTCA	235	
DB	158	ATGCGAGGATCGAC	CCCACTTTAAAGGTC	CTCTGTGACGTGAGTTCACCAACTTCTCA	217	
OY	236	TCACAGTTTACGCT	CGGTACCAAGACCA	CAATGAGACTACCGGCTGAATGCTCTTCTTGC	295	
DB	218	TCACAGGCTTGTCT	TATCGCGGAGCAAC	CAACATGAGACTACAGGTCACATCTCTCTGA	277	
OY	296	GGCAACAGTGGAA	TATGACCGCCTGTCT	TACCGAGAAATATCTGATAGCTCTTGAGCC	355	

D	b	278	GGCAGCAGTGGAAACGACCCCGCGTCGCGCTACATGAATATCCGACGACCTCTGGACC	337
O	y	356	TCGATCCCTCCATGCTGGACTCTATCTTGGAAAGCCAGACCTCTTTTGGTAAATGAGAAG	415
D	b	338	TTGACCCATCATCTTTGGATTCATCTGGAAAGCGCTGACTTGTCTTTTGGCCAAATGGAAG	397
O	y	416	GGGCGCACTTCATGAGGTGACCCAGCAACAAAGTTACTGCGCATCTTCAAGAAATGGGA	475
D	b	398	GGGGCCACTTCACAGACATCACCGACGACAAACAAAGCTGCTGAGAAATCTCCCGAAGCGCA	457
O	y	476	ATGTCGTGTACAGCATCAGGCTGACCGCTCATTTTGTCTGCTGATGAGACTCAGAACT	535
D	b	458	ACGTCCTTACAGCATGAGATCACCGCTGACTGTGGCTGCCATGGACCTGAAGAAATT	517
O	y	536	TCCCCATGACATCCAGACGTGACACGATGACAGCTTGAGAGCTCATCTCTGCAGCC	595
D	b	518	TCCCGATGTGATGATCAGACATGATATCAATGCAACTGGAAAGCT-----	559
O	y	596	CTGTCGCATCTGTGCACTTTCAGTGGCTCACCATGAAGAACCTGCTGTTGAGTGGC	655
D	b	560	-----TTGGTTATACCATGAACGACCTCATCTTTGATGTCG	595
O	y	656	TGGAAGATGCTCTGCTGTCCAAAGTGGCTGAGGGGCTGACTGTGCCCAAGTTATCTTGC	715
D	b	596	---AAGACCAAGAGAGCTGTGCAAGTGGCAGACGAGCTGACCTGCTCAGTTTATCTCGA	652
O	y	716	GGGATGAGAGAGATCTAGCGTGTGTTACCAAGCACTACACACAGGAAATTCACCTGCA	775
D	b	653	AGGAAGAAAGAAATCTGAGATACGTGACCAAGCACTACACACAGGTAATTCACCTGCA	712
O	y	776	TGGAAGTAAAGTTTCACTCGAGAGGGCAGATGGGCTACTATCTGATTCAGATGATACCTC	835
D	b	713	TTGAGGCCGATTCACCTGGAAGCGCAGATGGGCTACTACCTGATTCAGATGATACCTC	772
O	y	836	CCAGCTACTCATGCTCATCTCTGTCCTGGTCTCTCTTGATCAACATGAGATCTGCCC	895
D	b	773	CGAGCTCTTATGCTCATCTGTCGTCCTGATCTCTCTTGATCAACATGAGATCTGCAC	832
O	y	896	CTGCCGTGTGGGCTGGGCACTCACACCGCTGCTACCATGACCAACCAGAGCTGTGGCT	955
D	b	833	CAGCTCTGTGGGACTGTGGGCACTCACCAAGTCTACCATGACCAACAGAGCTGTGGCT	892
O	y	956	CCCGGGCCCTTTGGCTAAGGTGTCCTACGTGAAGGCAATGAGATCTGATGGCTGTGT	1015
D	b	893	CCCGAGCCCTCCCTACCCCAAGGTGCTCTACGTGAAGGCTATGACATCTGAGATGGCTGTT	952
O	y	1016	GTCCTCTTGTGTGTTGCTGCGCTTGTGTAAGATGCTGCATAAATTTTGTTCGCTC	1079
D	b	953	GGCTCTCTTGTGTGTTGCTGCGCGCTGTAAGATATGCGCGTGTCAACTTGTGTCCGGC	1012
O	y	1076	AGCATAAAGATTCATACGACTTGAAGAAGGACGAGCCGACACGCTTGGAGAAAGTA	1135
D	b	1013	AACCAAAAGAACTC-----CTTCGATTTAGAGAGAAAGCGGAGCATCACAAGGATGATG	1066
O	y	1136	TCATCCAGAAAGTGTGTTCTATTTCCGTGGGTATGGGTGGGGC---ACAGCCGACAGG	1192
D	b	1067	AGGTGGAGAAAGCGCTTCAACTTCTCGGCTATGGATGGGCGCCAGCCTGCTTGCAAG	1122
O	y	1193	CAGAAGTGGAGGTCCAAATGGAAGGTTCTGCAATTATAGTCCCAACCTCGAGCCCTC-	1251
D	b	1127	CCAAAGATGGCATCTCTGTCAAGGGTGGCAACAAACAAACGACGAAACCCCGCTCTCG	1186
O	y	1252	--CTTCTAAGGAGAGAGAAACACGCGGAAACTTACGTGAGTGTGACCAAGAAATTG	1305
D	b	1187	CACCGTCCAAAGTCCCGGAGGAGATGCGGAAACTCTTCACTCGAGAGACCAAGAGATCG	1246
O	y	1310	ACACCATCTCCGGGCTGTCTCCCTTTCACCTTCCATCTTCATCTCAATCTTCTAGCTGG	1365
D	b	1247	ACAAAATATCTGCAATCGGTTTCCCATGAGCCCTTCTCATCTTCAACATGTTCTACTGGA	1306
O	y	1370	TTGTCTAAATGCTATGTCAGAGAATATCACCA	1406
D	b	1307	TCATCTACAAGTTCGTCGGAGAGAGAGACGTCCACAA	1343

RESULT 12  
RATGLYRAL  
LOCUS  
DEFINITION  
Rattus norvegicus Glyralphal mRNA for glycine receptor alpha 1,  
complete cds.  
D00833.1 GI:220750  
ACCESSION  
D00833  
VERSION  
D00833.1  
KEYWORDS  
glycine receptor alpha 1; GlyRalphal.  
Rattus norvegicus (strain:Mistar) older than 25 days spinal cord  
cDNA to mRNA.  
ORGANISM  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
AUTHORS  
1 (sites)  
Grenningloh, G., Riesenitz, A., Schmitt, B., Methfessel, C., Zensen, M.,  
Bayreuther, K., Gundelfinger, E.D. and Betz, H.  
The strychnine-binding subunit of the glycine receptor shows  
homology with nicotinic acetylcholine receptors  
Nature 328 (6127), 215-220 (1987)  
JOURNLT  
MEDLINE  
87258250  
REFERENCE  
AUTHORS  
2 (bases 1 to 2125)  
Aragi, H., Hirai, K. and Hishinuma, F.  
Functional properties of strychnine-sensitive glycine receptors  
expressed in Xenopus oocytes injected with a single mRNA  
Neurosci. Res. 11 (1), 28-40 (1991)  
JOURNLT  
MEDLINE  
91367372  
REFERENCE  
AUTHORS  
3 (bases 1 to 2125)  
Hishinuma, F.  
Direct Submission  
Submitted (28-JUN-1991) Fumio Hishinuma, Mitsubishi Kasel Institute  
of Life Sciences, Department of Molecular Biology, 11 Minamiooya,  
Machida, Tokyo 194, Japan (Tel:0427-24-6233, Fax:0427-25-1252)  
FEATURES  
source  
1..2125  
/organism="Rattus norvegicus"  
/strain="Mistar"  
/db\_xref="taxon:10116"  
/tissue\_type="spinal cord"  
/dev\_stage="older than 25 days"  
/note="314 bp upstream of HindIII site"  
1..2125  
/gene="GlyRalphal"  
185..1534  
/gene="GlyRalphal"  
/codon\_start=1  
/product="glycine receptor alpha 1 precursor"  
/protein\_id="BAA0707.1"  
/db\_xref="GI:220751"  
/translation="MYSFNLRFLYMETIVEFSLAASKEDAAARSAPKPMSPDLK  
LMGRTSGDARIRNPNFSPVNSCNFINSPSIATETDQYINILRLOOMNDPRA  
YNEYRPSDLSDPSMDLSIKRDLFPANEGARHETTTNKLIRISNGNVLSITL  
TLTACPMDLKRFPMVDVQTCIMQLESFYTMNDLIFEMODGAVQADGLTLPQTLK  
EERDLRYCTNHYNTGKFTCLFARHLERONGYLIOMYIPSLIVLSLWISFWINMDA  
APAVVGITITVLTMTQSSGSRSALSKPVSYKVIDIMAVVCLIFVSALLEYAAVNF  
VSROHKLLEFRKRHRHKKDEGEGEFNSAYGAPACLOAKGIVSGANNNTN  
PAPAPSKSPKMRKLFIQRAKKIKDKISRGIPMAFLINNEYWIIYKIVREDVHNK"  
185..268  
/gene="GlyRalphal"  
/gene="GlyRalphal"  
269..1531  
/product="glycine receptor alpha 1"  
BASE COUNT 536 a 593 c 527 g 469 t  
ORIGIN  
Query Match 39.9%; Score 654.6; DB 10; Length 2125;  
Best Local Similarity 72.6%; Pred. No. 6.6e-171;  
Matches 941; Conservative 0; Mismatches 299; Indels 57; Gaps 5;  
OY 116 CCCAGCCATGTCCTCGATTTCTAGACAACATTATGGCGGCAACATCTGATATG 175  
II IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII IIII

Db 282 CCAAGCCTATGTCACCTCGACCTCTGGATTAAGCTTATGGAGAGACTTCTGGATG 341  
OY 176 ATGCCAGGATTCGGCCCAATTTAAAGCCACCACCGAAGCAGTACCTGCAACATCTCA 235  
Db 342 ATGCCAGGATTCAGCCCAACCTTAAAGTCTCTCTGGAACGTAAGTGAACATCTTCA 401  
OY 236 TCAACAGTTTCAGCTCCGTCACCAAGACCAACAAATGAGACTCCGGGTAAATGCTTTCG 295  
Db 402 TCAACAGCTTTGGTTCATATGCGCCAGACCAACATGAGCTACAGGGTCAACATCTTCTGA 461  
OY 296 GGCACAGTGAATGATACCCAGCCCTGCTACCGAGAAATATCTGATGACTCTCTGGACC 355  
Db 462 GGCAGCAGTGAAGAGACCCCGCTCTCGCTACATGAATATCTGAGAGACTCTCTGGACC 521  
OY 356 TGCATCCCTCATGCTGACCTCATATGGAAGCAGACCCCTCTTGTGTAATGAGAAG 415  
Db 522 TTGACCATCATGTTGATTTCCATCTGGAAGCCCTGCTTTTGTGCAATGAGAAG 581  
OY 416 GGGCCAACTTCATGAGTGCACGACGACAAACAAGTACTGCGCATCTTCAAGAAATGGA 475  
Db 582 GGGCCCACTTCACAGAGATCACACGACAAACAAGCTGTGAGAAATCTCCGGAGCGCA 641  
OY 476 ATGTGCTGTACAGCATAGGCTGACCTCATTTTGTCTGCTGATGAGACTCAAGACT 535  
Db 642 ACCTGCTGTACAGCATAGGATCACCTGACCTGCGCTCCCATGAGACTCAAGAAAT 701  
OY 536 TCCCATGAGACATCAGACGACGACGATGAGCTGAGAGCTCATCATCTGACAGCC 595  
Db 702 TCCCATGAGACATCAGACATGTAATGACACTGGAAGCT----- 743  
OY 596 CTGTGCATCTGTGTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 655  
Db 744 -----TTGGTTAATACATGAAGACAGCTCATCTTGTGAGTGGC 779  
OY 656 TGAAGATGCTCTGCTGTCCAAAGTGGCTGAGGGGCTGACTGTGCCAGTTATCTTGC 715  
Db 780 ---AAGAGCAAGAGAGCTGTGAGTGCAGATGAGTGCAGTCCCTCCAGTTATCTCTGA 836  
OY 716 GGGATGAGAGAGATCTAGGCTGTGTACCAAGCACTCAACACAGGAAATTCACCTGCA 775  
Db 837 AGGAAGAGAGAGATCTAGATACCTGACACCAAGCACTCAACACAGGTAATTCACCTGCA 896  
OY 776 TCGAGTAAAGTTTCACTGAGAACGAGATGGGCTACTATCTGATTCAGATGATATCC 835  
Db 897 TTGAGGCCCATTTCCACTGGAAGCGAGATGGGCTACTACATGATCCAGATGATATCC 956  
OY 836 CCAGCCTACTCATGATGATCTCTGCTGCGGCTCTCTCTGATCAACATGATGATGCTGCC 895  
Db 957 CCAGCCTGCTTATGTCATCTCTGCTGAGATCTCTCTGATCAACATGATGATGCTGCC 1016  
OY 896 CTGCGCCTGAGGCTGGGATCCACACCGTGTCTCACATGACACCCAGAGCTCTGGCT 955  
Db 1017 CAGCTGCTGAGGATCGGATCGGATCCACACAGTGTCTCACATGACACCCAGAGCTCTGGCT 1076  
OY 956 CCGGCGCTCTTGTGCTTAAGGTGTCTACGTGAAGCAATGACATGATGATGCTGTGT 1015  
Db 1077 CCGGCGCTCTCTTCCATCCCAAGGTGTCTACGTGAAGCAATGATGATGATGCTGTGT 1136  
OY 1016 GTCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1075  
Db 1137 GCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1196  
OY 1076 AGCATTAAGAAATGATACGATTCGAGAAAGAGAGAGAGAGAGAGAGAGAGAGATG 1135  
Db 1197 AACACAGATG 1250  
OY 1136 TCATCAAGAAAGTGTGTTCTATTTCCGTGCTATGCTTGGCTTGGCTTGGCTTGGCT 1250  
Db 1251 AGGTGATG 1250  
OY 1193 CAGAGATGATG 1250  
Db 1311 CCAAGGATGAGATGATG 1311



RESULT 13  
HSLYRA1 1857 bp mRNA linear PRI 28-MAY-1993  
DEFINITION H.sapiens alpha-2 strychnine binding subunit of inhibitory glycine receptor mRNA.  
ACCESSION X52008.1 GI:31848  
VERSION 1  
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1857)  
Greenan, J.G., Schmieden, V., Schofield, P.R., Seeburg, P.H., Siddique, T.K., Mohandas, T.K., Becker, C.M. and Betz, H.  
Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes  
EMBO J. 9 (3), 771-776 (1990)  
JOURNAL MEDLINE 90183975  
PUBMED 2155780  
FEATURES  
Source location/Qualifiers  
1.1857  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
/dev\_stage="fetal"  
388..1746  
/gene="GLYRA1"  
388..1746  
/gene="GLYRA1"  
/note="strychnine binding alpha-2 subunit"  
/codon\_start=1  
/product="inhibitory glycine receptor"  
/protein\_id="CAA36257.1"  
/db\_xref="GI:31848"  
/db\_xref="SWISS-PROT:P23416"  
/translation="MNRLVNLITLAFPLETNHFTAFCKDHSRSGKQPSQTLSP  
SDFLKMGRTSGYDARIRPNFKGPVAVTNIINFSFGSTETMDYRNIPLRQOM  
NDSRLAYSEYDDSLDIPMSDLSIMKDLFEANKEKGNFHDVTDNKLRSKGV  
LXSLRLTLTSCPMIDKMPMDVOCITMOLSFQYTMNDLFEMLSDPQVABSLTL  
PQFLIKERKEIGCTKHYNTKFTICIEKPHLEKMGYLLIOMTIPSLIYILSVSF  
WIFINDAARVALGLITVLTMTQSSGSRADLPKYSYKALIDIMAVCLVFALALE  
YAAVVFVSRQKELRLRRQKRONKEDVTRSFNSGVGMGCLQKGTAVKAT  
PANLPQPPKDGDAIKKFKVDRAKRIDTISRAPFLAFLINIFWITIKIRHEDVH  
KK"  
BASE COUNT 526 a 416 c 417 g 498 t  
ORIGIN  
Query Match 39.6%; Score 650.2; DB 9; Length 1857;  
Best Local Similarity 70.1%; Pred. No. 1.1e-169;  
Matches 943; Conservative 0; Mismatches 348; Indels 54; Gaps 3;  
QY 63 CCTCTCAGAGGTGCGCTTGCGCAAGAGAGCAATGTCAGACCAAGGGTCCAGCC 122  
DB 450 CTTTCAGAGCGCTTTCGCAAGAGCAATGTCAGAGTCTGGAACAACCTTCACAGC 509  
QY 123 CATGTCCTCCCTGATTTCTGACAACTTATGGGCGACATGTCGATGATGATGCGCAG 182

DB 510 CTTATCTCTTCAGATTTCTTGACCAAGTATATGGAGAG  
QY 183 GATTGGCCCAATTTTAAAGGCCACCCGTGAACGTGACC  
DB 570 AATCAGGCCAAATTTTAAAGGTCCTCCAGTAAGAGTTACTTGG  
QY 243 TTTCAGCTCCGTACACCAAGCCAAATGAGACTACCGGGTGAAAT  
DB 630 TTTTGGATCAGTACACAGAAAGACCATGAGACTACCGGATGAAT  
QY 303 GTGGAATGACCCACGCGCTGCTACCGAGAAATTCGATGACTCT  
DB 690 GTGGAATGATTCACGCGCTGGCGTACAGTATGACCGAGTACGCTCC  
QY 363 CTCATGCTGCTGACTATATGGAAGCCAGACCTCTTTGTAATGAG..AAGGGCCAA 422  
DB 750 ATCCATGCTAGACTACATTTGGAACAGATTTGTTGTCCTTGGCAATGGAAGGTGCA 809  
QY 423 CTTTCATGAGGTGACCAAGGACCAAGTATGAGTCCGATCTTCAAGAAATGGGAATGCT 482  
DB 810 CTTCCAGAGTGTACACCTGACCAACAAATGCTACGGAATTCGAAATGGAAGTGTCT 869  
QY 483 GTACAGATGACGCTGACCCCTCATTTTGTCTGCTGATGAGCTCAGCAACTTCCCAT 542  
DB 870 CTACAGATGACGACTACACCTTACCTTATCCTGTCCATGAGTGAAGAACTTCCGAT 929  
QY 543 GGACATCCAGACGTGACAGATGACAGCTTGAAGCTCATCTGACAGCCCTCTGCC 602  
DB 930 GGATGTCACAGACCTGTACAAATGACGTGAGAGTT..... 964  
QY 603 ATCTCTGTCACCTTTCAGTGGCTACACCATGAAGACCTGCTGTTGAGTGGCTGAAGA 662  
DB 965 -----TTGGGTACAGGATGAATGACTGATATTTGATGAGTGAAGTGA 1007  
QY 663 TGCTCTGCTGCTGCCAGTGGCTGAGGGGCTGACTCTGCCAGTTTATCTTGGCGATGA 722  
DB 1008 TGTGTC---CACTGCAAGTGTGCTGAAGATGTACCTCGCCAGTTTATTTGAAGAAGA 1064  
QY 723 GAAGATGATGAGCTGTGTTACCAAGCACTACACACAGGGAATTCACCTGATCGAGGT 782  
DB 1065 GAAGGACCTTGGCTACGTGTACAAAGCACTACAAACATGTGAAGGTTTACCTGATTTGAGT 1124  
QY 783 AAGATTCACCTGGAAGGACAGATGAGGCTACTATCTGATTCAGATGATATCCCAAGCT 842  
DB 1125 CAGATTTACCTGGAAGGCCCAATGAGATATTTGATTCAGATGATATCCCAAGCT 1184  
QY 843 ACTATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 902  
DB 1185 GCTTATGATTAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1244  
QY 903 TGTGAGGCTGGGCACTACACCCGCTGCTACCATGACCAAGGAGCTGCTGCTGCTGCTG 962  
DB 1245 GGTGCACTGGGCACTACACAGTCTTACCATGACCAAGGAGCTGCTGCTGCTGCTGCTG 1304  
QY 963 CTTCTTGGCTAAGGTGCTCTACGTGAAGGACATGACATGAGTGGTGTGTGCTGCT 1022  
DB 1305 ATCTTGCCAAAGGCTCTCTATGTAAGGATGACATTCGTGATGAGTGGGCTGCTGCT 1364  
QY 1023 CTTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1082  
DB 1365 GTTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1424  
QY 1083 AGAATTCATGACACTTGCAGCAAGGACCAAGGCGCCCAAGCTTGGAGGAAGATATCATCA 1142  
DB 1425 GGAGTTTCTGCGCCCTCCCAAGGAAGCAAGAGGAGATGAAGGAAGAGAGCTTACTGCG 1484  
QY 1143 AGAAGTGGTTTCTATTTCCGCTATGAGCTTGGGCACTGCTGCGAGGCAAGAGATGG 1202  
DB 1485 TGAAGTGGTTTATTTAGCGTTATGAGGATGAGTGGCTGCTGCGAAGGAAGAGATGG 1544  
QY 1203 AGTCCCAATGGAAGTCTTGCGATTTATGATCCCAACCTCCAGCCCTCTTTTGAAGGA 1262  
DB 1545 AACAGCTGTCAAGGCGCAC-----ACGTGCCAACCACTCCCAACAGCCCAAAAGA 1595



QY	1263	AGGAAACCCAGGGGAAACCTCTACGGGATGAGCCAAAGAAATTCACACATCTCCG	13322
DB	1396	TGGAGATGCTATCAAGAAAGATTGTGGACCGGGCAAAAAGATGACAGATATTCG	16555
QY	1323	GGCTGCTCTCCCTTTCACCTTCCATCTCCATCAATATCTTCTACTGGGTTGCTATMAAGT	13823
DB	1556	AGCTGCGCTCCATTTGGCGCTTCCTCATTTTTCACATCTTTTACTACGATACATACAGAT	17155
QY	1383	GCTATGGCTCAGAAATATCCACCG	1407
DB	1716	CATTCGGCATGAAAGATGTCACAAAG	1740
RESULT 14			
LOCUS	AF268375	1744 bp	mRNA linear
DEFINITION	Bos taurus glycine receptor alpha 1 subunit (Glr1) mRNA, complete cds, alternatively spliced.		
ACCESSION	AF268375		
VERSION	AF268375.1	GI:10180958	
KEYWORDS			
SOURCE	Bos taurus.		
ORGANISM	Bos taurus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovine; Bos.		
AUTHORS	1 (bases 1 to 1744)		
TITLE	Pierce,K.D., Handford,C.A., Morris,R., Vafa,B., Dennis,J.A., Healy,P.J., and Schofield,P.R.		
JOURNAL	A nonsense mutation in the alpha1 subunit of the inhibitory glycine receptor associated with bovine myoclonus		
MEDLINE	Mol. Cell. Neurosci. 17 (2), 354-363 (2001)		
PMID	21109390		
REFERENCE	11178872		
AUTHORS	2 (bases 1 to 1744)		
TITLE	Pierce,K.D., Handford,C.A., Morris,R., Vafa,B., Dennis,J.A., Healy,P.J., and Schofield,P.R.		
JOURNAL	Direct Submission		
FEATURES	Submitted (17-MAY-2000) Neurobiology Program, Garvan Institute of Medical Research, 384 Victoria Street, Sydney, NSW 2010, Australia		
SOURCE	Location/Qualifiers		
gene	1. 1744		
CDS	/organism="Bos taurus" /db_xref="taxon:9913" 1. 1744 /gene="Glr1" 229..1602 /gene="Glr1" /note="Glr1s; alternatively spliced" /codon_start=1 /product="glycine receptor alpha 1 subunit" /protein_id="AA014346.1" /db_xref="GI:10180958" /translation="MYSEFNTLRILYMETIVFFSLAASKAEARASKPMSPDFLDR LMTSGYDARIENFGKPPVAVSCNFIENFGSIAETMYRNVILROOQNDPRILA YNEYPDSDLDSPMDLSITKPDLEFENEGAHHEITDOKRLIRSRNGVLYSITG TTLTACPMDLKPNPDYOTCIMOLESQGYKNDLIFEMOEGAVQVADGLPLPILAK EERLIRCTKRYHNTGKTCIARFHLEROMKTYLIOMTIBSLVILVSWISFWIMDD APAVAGELITVTLVMTTQSSGRASLRKSVYKVIDIMAVCLFVSALLEVAVNVE VSROHKEILRRRRRHHKSPMLNLPQDEAGEGRFSYSAAGMPACLOAKDGSIVGK ANNSNTNPPRAPSPEEMKRLFIORAKKIDKISRIGFPWAFIFNFWIYIIVR REDVHNO"		
BASE COUNT	396 a 508 c 451 g 389 t		
Query Match	39.4%; Score 646.8; DB 4; Length 1744;		
Best Local Similarity	71.2%; Pred. No. 9.6e-169;		
Matches	956; Conservative 0; Mismatches 317; Indels 69; Gaps 5;		
QY	89	AGGAAGTCAATCTCGAAGCAAGGAGGTCCAGCCCATGTCCCCCTCTGATTTCTTACAGCA	148
DB	299	AGGAGGCTGAAGCTGCTCGCTCTTCAAGCCCATGTCAACGCTCCGATTTCTTGGA	358

QY	149	AACTTAATGGGGGGAACATCTGGATATATATGACAGATTCTGGCCCAATTTTAAAGGCCAC	208
Db	359	AACCTATATGGGAGGACCTTCTGGATATGACAGCCAGGATCAGGCCCAATTTCAAAGGTCTCC	418
QY	209	CCGTGAACGTACACCGCAACATCTTCAACACAGTTTACGCTCCGTCACCAAGACACAA	268
Db	419	CAGTGAATGTGACGTCGAACAATTTTCATTCACAGCTTGCGMTTCATTTGCTGACACACTA	478
QY	269	TGGACTACCGGGTGAATGTCTTCTTGGCGGCAACAGTGAATGACCACAGCCTGTCTTACC	328
Db	479	TGGACTACAGGGGTCAACATCTTCCGCGGCAAGCACTGGAAATGACCCCGGGTGGCTTACA	538
QY	329	GAGAAATATCTGATGATGATCTGTGACATCTGACATCCCTCATGCTGAGACTCTTATCTGAAGC	388
Db	539	ATGAATACCCCGATGATCTGCTGAGACTTGGACCATTCATGTGGACTCTCATCTGGAAAC	598
QY	389	CAGACCTCTTCTTGTCTAATGAGAAAGGGGCCAACCTTCATGAGGTGACACGAGCAACA	448
Db	599	CTGACCTCTTCTTGTCCAAATGAGAAGGGGGCCCACTTCATGAGATCAACAGACACACA	658
QY	449	AGTTACCTGGCATCTTCAAGAAATGGGAATGTCTGTACAGATACAGCTGACCCCTCATTT	508
Db	659	AACCTGCTAGGATCTCCCGAAGGGGAATGTCTCTTACAGATTAAGATCAACCTCGACAT	718
QY	509	TGTCCTGCTGATGAGACCTCAAGAACTTCCCATGAGATGACATCGACATGACAGTACAGC	568
Db	719	TGGCCTGGCCCATGATGACTTGAAGAAATTTCCCATGATGTCCAGACATGTCAATTAAC	778
QY	569	TTGAGAGCTCATCCATCTGTGACAGCCCTGTGCATCTGTCTACCTTTCAGTTGGCTACA	628
Db	779	TGGAAGTCT-----TTGGATACA	796
QY	629	CCATGAAGACCTCTGCTTTGAGTGGCTGGAAGATGCTCTGCTGCCAAGTGGCTGAG	688
Db	797	CCATGAAGACCTCATCTTTGAGTGGCAAGACAG-----GGGCTGTGCAAGTGGCAAGT	853
QY	689	GGCTGACTGTGCCCATGTTTATCTTGGCGGATGAGAGAGATCTAGGCTGTGTACCAAGC	748
Db	854	GACTTAACCTCTACCCGAGTTTATCTCGAAGAGAGAGAGAGACTTGAATCTGACACCAAGC	913
QY	749	ACTACACACAGGGAATTCACCTCTCATGAGAGTAAAGTTTACCTGGACGGCGAGATGG	808
Db	914	ATTACACACAGGCAATTCACCTGATCTGAGAGCCCGGTTCCACCTGGAGCAGACAGATGG	973
QY	809	GCTACTATCTGATTCAGATGTATACATCCCGAGCTACTCATGCTATCTCTGCTGGCTCT	868
Db	974	GCTACTACTGTATCCAGATGTATACATCCCGAGCTCTCATTTGTATCTCTGTGTGATCT	1033
QY	869	CCTTCTGATCAACATGATGATGCTGCCCTGGCCCTGTGTGGGCTGGGCAATCACACCGTGC	928
Db	1034	CCCTTCTGATCAACATGATGATGCTGTCCCGCCCGTGTGGGCTGGGCAATCACACCTGTGC	1093
QY	929	TCACCATGACACACAGAGCTGTGGCTCCCGGAGCTCTTGGCTTAAGGTGTCCTACGTGA	988
Db	1094	TCACCATGACACACAGAGCTGTGGGCTCTCGAGCATCTCCCTGCCCAAGGTATGTATGTGA	1155
QY	989	AGGCAATGACATCTGATGATGAGCTGTGTCTGTCTCTTGTGTCTGCTCTCTGTGAGT	1048
Db	1154	AGGCAATGACATCTGATGATGAGCTGTGTCTGTCTCTTGTGTCTGACACCGCTTGTGAGT	1213
QY	1049	ATGCTGCATTAATTTTGTCTGTCTGTCAGCAATAAAGATTTCAATACGACTTGGAGCAAGC	1108
Db	1214	ACGCGCGGCTCAACTTGTGTGTGTCGGCAGCAAAAGAGACTGTCTCCGATTCACAGAGGAAGC	1277
QY	1109	AGAGCGGCC-----AACGCTTGAGGAAGATATTCACCAAGAAAGTCT	1150
Db	1274	GAAGCATACACAAGAGCCCATGTGTGATATGTGTTCACAGAGAGATGAGCTGGAAGAGGCGC	1333
QY	1151	GTTTCTATTTTCCGTGCTATAGGCTTGGGCG---ACTGGCTGACGGCAAGACATGAGAGTCT	1207
Db	1334	GCTTCAACTCTTCTGGCGCTAAGGCAATGGGCGCCAGCCTGTCTTACAGGCCAAGAGTGGCATCT	1393

QY 1208 CAATGAGGTC---TGCAATTATAGTCCCAACCTCCAGCCCTCTTCTAAGGAG 1264  
Db 1394 CTGTCAAGGCGGCAACACGACGACCAACCTCTCCGGGACCATCCAGTCC 1453  
QY 1265 GAGAAACACGCGGAACTCTAGTGAGGACGACCAAGATGACCATCTCCGGG 1324  
Db 1454 CGAGAGAGATGCCAAAACCTTTATCCAGCGGCGACGAGATGACAGATCTCCGCA 1513  
QY 1325 CTGTCTCCCTTCACTTCCATCTTCAATATCTTCTACTGGGTGTCTATAAGTGC 1384  
Db 1514 TCGGCTTCCCATGGCTTCTCATCTTCAACATGTTCTACTGATCATCTACAAAATCG 1573  
QY 1385 TATGCTCAGAGATATCCACCA 1406  
Db 1574 TCCGAGGAGGAGCTCCACCA 1595

RESULT 15  
LOCUS S73718 1431 bp mRNA linear ROD 17-JUL-2000  
DEFINITION Glra1=Inhibitory glycine receptor alpha 1 subunit [alternatively spliced] (mice, A/HeJ, brainstem, spinal cord, mRNA Mutant, 1431 nt).  
ACCESSION S73718  
VERSION S73718.1 GI:765208  
KEYWORDS  
SOURCE Mus sp.  
ORGANISM Mus sp.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1431)  
Ryan, S.G., Buckwalter, M.S., Lynch, J.W., Handford, C.A., Segura, L., Shiang, R., Wasmuth, J.J., Camper, S.A., Schofield, P. and O'Connell, P. A missense mutation in the gene encoding the alpha 1 subunit of the inhibitory glycine receptor in the spasmodic mouse  
Nat. Genet. 7 (2), 131-135 (1994)

JOURNAL MEDLINE  
PUBMED 9504575  
REMARK 7920629  
Genbank staff at the National Library of Medicine created this entry [NCBI g1b5g 156497] from the original journal article.  
This sequence comes from Fig. 2a and 2b.  
Location/Qualifiers  
1..1431  
/organism="Mus sp."  
/db\_xref="taxon:10095"  
/tissue\_type="brainstem and spinal cord"  
/note="A/HeJ"  
1..1431  
/gene="Glra1"  
58..1431  
/gene="Glra1"  
/note="alternatively spliced; this sequence comes from Fig. 2a and 2b; conceptual translation differs from that in published reference; Glra1 protein; Glyc alpha 1"  
/codon\_start=1  
/product="inhibitory glycine receptor alpha 1 subunit long form"  
/protein\_id="AAB32158.2"  
/db\_xref="GI:9247216"  
/translation="MYSFNTLRFLYMETIVFESLASKEAEARAPKPMSPDFLDK LMGRTGYDARIPNKGPPVANSYCNIFINSFISIAETMDRYNVLFRQOMNDPLA VNEYPPDSLDLPMSMDITWKDLPFRANKGAHFHETTDNKLIRSRGNVLSIRI TITLACPMILKNEPMDVOTCIQMLESEFGTMDLIREWDEQGVAVADGLTLPQFLK EKKDLRYCTKHNTGKFTICIEARFILERQMYLLQMYLPSLIVLSISFMINDA APARVGLITTVLTMTTOSGSRASLPKYSYKVIDIMAVCLLEFVSLIETAAVNF VROHKELEFRKRRHRSKSPMLNLFODEGEGRNFSAYGMPACLOAKGDISVG ANNNNTNPPAPSPKSPKPEMRKLFIDIRAKIKDISIRGPMFLINMEYIIYIKIVR REDVHNK"

BASE COUNT 355 a 411 c 336 g 329 t  
ORIGIN

Query Match 39.4%; Score 645.4; DB 10; Length 1431;  
Best Local Similarity 71.9%; Pred. No. 2.3e-168;

Matches 945; Conservative 0; Mismatches 301; Indels 69; Gaps 5;  
QY 116 CCCAGCCATGTCCTCCCTGATTTCTTACACAACCTTATGGGGCAACATCGATATG 175  
Db 155 CCAAGCCTATGTACCTCTGGACTTCTCTGATTAAGCTCATGGGAGACTTCTGGGTATG 214  
QY 176 ATGCAAGATTGCGCCCAATTTTAAAGGCCACCGGTGAACGTAAGCTGCAACATCTTGA 235  
Db 215 ACCCAGAGATCAGGCCCAACTTTAAAGTCTCTGTAAGCTTAAGTTGCAACATCTTCA 274  
QY 236 TCAACAGTTTCACTCGGTGACCAAGACCAATGACACTACCGGGTGAATGTTCTTTCG 295  
Db 275 TCAACAGTTTGGTTCATCTGATGAGACCAACCATGACATATAGGTAACATCTTCTTGA 334  
QY 296 GGCACAGTGAATGAGCCCAAGCCCTGCTACCGAGAAATATCTGATGACTCTGGACC 355  
Db 335 GGCAGAGTGAAGACGACCCCGCTGGCTTACATATAATACCTGATGACTCTTGACC 394  
QY 356 TCGATCCCTCATGCTGAGACTTATCTGNAAGCCAGACCTCTTGTGTAATGAGAAG 415  
Db 395 TTGACCATCTATGTTGATTCATCTGGAAGCTGACTGTTCTTGGCAATGAGAAG 454  
QY 416 GGGCCCACTTCCATGAGTGACGACGACACCAAGTACTGGCATCTTCAAGATGGA 475  
Db 455 GGGCCCACTTCCCAAGTAATCACCGACGACCAAACTGCTAAGAACTCCCGGAATGGA 514  
QY 476 ATGTGCTGTACAGCATGAGCTGACCTCATTTTGTCTGCTGATGAGACTTCAAGACT 535  
Db 515 ATGTGCTGTACAGCATGAGACTGACCTCATTTTGTCTGCTGATGAGACTTCAAGACT 574  
QY 536 TCCCATGAGCATTCAGAGCTGACAGTGCAGATGCAAGCTTGAAGCTCATCTCTGAGCC 595  
Db 575 TCCCATGAGTGTACAGAGCTGATATCATCAACTCCAAAGCT----- 616  
QY 596 CTGTGCATCTGTGCTACTTTCAGTTGCTGACACCAATGAAAGCTCTGTTGAGTGGC 655  
Db 617 -----TTGATATACCAAGACGACTCATCTTGAATGGC 652  
QY 656 TGGAGATGCTCTGCTGTCACAGTGGCTGAGGGGCTGACTGCCCCAGTTATCTTGC 715  
Db 653 AGAGAGAG---GAGTGTGAGAGTGGGACGAGAGACTGACCTGCTCATTTATCTGA 709  
QY 716 GGGATGAGAGATCTAGGCTGTGTACCAAGCATACACACAGGGAATTCACCTGCA 775  
Db 710 AGGAAAGAAAGACCTGAGATCTGCACCAAGCATACACACAGTAATTCACCTGCA 769  
QY 776 TCGAGTAAAGTTTCACTGGAAGGAGGACAGTGGGCTGATGATGATGATGATGATGATG 835  
Db 770 TCGAGGCGCGATTCACCTGAGAGCGGAGATGGGCTACTACCTGATCCAGATGTACATCC 829  
QY 836 CCAGCTACTCATCTCATCTGTCTGCTGCTCTCTTCTGATCAACATGATGCTGCC 895  
Db 830 CCAGCTGCTCATCTGATCTGCTGATCTCTGATCTCTTCTGATCAACATGATGCTGCC 889  
QY 896 CTGCGCTGTGGGCTGGGCTGATCAGACCTGCTTACCATGACCAACCCAGAGCTGGCT 955  
Db 890 CAGCTGTGTGGGCTGGGCTGATCAGACCTGCTTACCATGACCAACCCAGAGCTGGCT 949  
QY 956 CCGGCGCTCTTCTGCTAAGGTTCTCTAGCTGGAAGCATGCACTGATGAGGCTGTCT 1015  
Db 950 CCGGAGCTCTCTCAACCAAGGTTCTCTAGCTGGAAGCATGCACTGATGAGGCTGTCT 1009  
QY 1016 GTCTGCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075  
Db 1010 GCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1069  
QY 1076 AGCATTAAGATTCATACGATCTGAG-----AAGCAGAGGCGCC 1117  
Db 1070 AACACAAGAGAGCTTTCGATTTAGAGAGAGGAGGAGACATCAACAAGCCCATGCTAA 1129  
QY 1118 AACGCTTGAGAGATATCATCAAGAAGTGTCTTATTTCCGTGCTTATGCTTGG 1177  
Db 1130 ATCTGTTTCAAGGATGAGAGGTTGAGAGGCGGCTTCAACTTCTCTGCTATGGATGG 1189



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 22:23:50 ; Search time 3552.13 Seconds

(without alignments)  
3531.204 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTTLPVATLSFLMLTLPLGQ.....PQPPAPLREGETRKLYVD 431

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPRO/US10075846/runat\_25062003.163648.5147/app.query.fasta\_1.782  
-DB=genembl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846.ecgn.1.1.384.4.runat.25062003.163648.5147 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGOQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=200 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: GenEmbl:\*  
2: gb\_hkg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_jnu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hkg\_hum:\*  
31: em\_hkg\_jnu:\*  
32: em\_hkg\_other:\*  
33: em\_hkg\_mus:\*  
34: em\_hkg\_pln:\*  
35: em\_hkg\_fod:\*  
36: em\_hkg\_mam:\*  
37: em\_hkg\_vrt:\*  
38: em\_sy:\*  
39: em\_hkg\_hum:\*  
40: em\_hkg\_mus:\*  
41: em\_hkg\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2163	96.0	1874	6 AX392950	AX392950 Sequence
2	2157	95.8	1600	6 AX037565	AX037565 Sequence
3	1886	83.7	1251	10 AF462147	AF462147 Mus muscu
4	1749.5	77.7	2045	5 DRE404970	AJ404970 Danio rer
5	1682.5	74.7	1857	9 HSGLYRA1	X52008 H.sapiens a
6	1677	74.5	1359	10 RNO310837	AJ310837 Rattus no
7	1677	74.5	3160	10 RNIGRAA2	X61159 R.norvegicu
8	1677	74.5	3865	10 RNNEGLY	X57281 Rat NGlyR m
9	1609	71.4	3069	9 HSU93917	U93917 Human glycl
10	1598	71.0	1443	10 RNO310838	AJ310838 Rattus no
11	1592	70.7	2413	10 RATIGRAS	M55250 Rat inhibi
12	1580.5	70.2	1715	9 HSGLYRA2	X52009 H.sapiens a
13	1580	70.2	1744	4 AF268375	AF268375 Bos tauru
14	1578.5	70.1	1392	10 AF362764	AF362764 Mus muscu
15	1577	70.0	1628	5 DRE5812	AJ005812 Danio rer
16	1570.5	69.7	1407	10 S73717	S73717 GIRA1+Inhib
17	1569.5	69.7	1350	10 RNO310834	AJ310834 Rattus no
18	1569.5	69.7	1431	10 S73718	S73718 GIRA1+Inhib
19	1569.5	69.7	2125	10 RATGLYRA1	DO0833 Rattus norv
20	1568.5	69.6	1374	10 RNO310835	AJ310835 Rattus no
21	1568.5	69.6	1679	10 RNIGRA1	X55246 R.norvegicu
22	1564.5	69.5	1384	10 RNRGRSTR	Y00276 R.norvegicu
23	1559.5	69.2	1682	5 AF488379	AF488379 Danio rer
24	1554	69.0	1418	5 AT094574	AT094574 Morone am
25	1537.5	68.3	2376	5 DRE308516	AJ094975 Morone am
26	1494.5	66.4	3085	5 AY094975	AY094975 Morone am
27	1413	62.7	1222	10 RNO310836	AJ310836 Rattus no
28	1381	61.3	1595	5 DRE308517	AJ308517 Danio rer
29	983.5	43.7	657	5 E03608	E03608 DNA encodin
30	876.5	38.9	2160	5 DRE404971	AJ404971 Danio rer
31	871	38.7	2404	6 AX305334	AX305334 Sequence
32	871	38.7	2404	6 MU093939	U09399 Mus musculu
33	863.5	38.3	1491	10 RNO310839	AJ310839 Rattus no
34	858.5	38.1	1465	10 MMGRBMA	X81202 M.musculu
35	855.5	38.0	1737	9 AF094754	AF094754 Homo sapi
36	855.5	38.0	1737	9 AF094755	AF094755 Homo sapi
37	855.5	38.0	2106	9 H5033267	U33267 Human glycl
38	855.5	38.0	2649	9 BC032635	BC032635 Homo sapi
39	850.5	37.8	1661	4 AF268376	AF268376 Bos tauru
40	761	33.8	39796	6 AX037570	AX037570 Sequence
41	761	33.8	39796	6 HSU3563	Z93848 Human DNA s
42	757	33.6	1446	6 AX477445	AX477445 Sequence
43	757	33.6	1446	6 AX477446	AX477446 Sequence
44	757	33.6	5503	6 AX477442	AX477442 Sequence
45	757	33.6	5503	6 AX477444	AX477444 Sequence

RESULT 1

## ALIGNMENTS





CDS

1. 1251  
/gene="Glyr4"  
/note="ligand-gated ion channel"  
/codon\_start=1  
/product="glycine receptor alpha 4 subunit"  
/protein\_id="AAL69899.1"  
/db\_xref="GI:18448711"  
/translation="MSPSDFLDKMGRTSGYDARIKPNKGPVNVTCNIFINSEGSV  
TETMDYRVNVDKQNDPRLAYREYDDSLDNLNPSMLESIMKPDLEFANEKANFH  
EETDNDKLRIFKNGNVLISRLILISCPMDLKNFMDIOTCMQLESFGYTNMDLM  
FEMLEDAPOVAVAGLTIPORILDEKDLGYCTKHYNPKFTEIEVKHLEKMGYYL  
IQMTPLSLIYILSVSFNIMMDAPARVAGITTYVLTMTQSSSRASLEKSVYKA  
IDIMAVCLLEFVALLLEYAANFVSROHKEFMRRLRRORRMEEDIRSRFTFRG  
YLGHCLOARDGGMESGIYSQPPTPLKEGETMRLKYVDRAKRIDITISRAVFPPT  
FLVNFIFWVYVKLRSEDHIAL"

BASE COUNT 324 a 314 c 281 g 332 t

ORIGIN

Alignment Scores:

Pred. No.: 6,05e-182 Length: 1251  
Score: 1886.00 Matches: 358  
Percent Similarity: 94.36% Conservative: 10  
Best Local Similarity: 91.79% Mismatches: 8  
Query Match: 83.75% Indels: 14  
DB: 10 Gaps: 1

US-10-075-846-4 (1-431) x AF462147 (1-1251)

OY 42 MetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArg 61  
DB 1 ATGTCCCTCTGATTTCTTGACAAAGCTTATGGAGCAACATCTGGATATGATCTAG 60  
OY 62 IleArgProAsnPhelYsglyProProValAsnValThrCysAsnIlePheIleAsnSer 81  
DB 61 ATTAGGCCCAATTCAAAGTCCACCTGATGTGACCTGCACATCATCTTCAACAGT 120  
OY 82 PheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGln 101  
DB 121 TTTGGCTCGCTACGAGACCAACCATGGACGTACGGGTAATGTCTTCTTGGCGACGCG 180  
OY 102 TrpAsnAspProArgLeuSerTyrArgGluTyrProAspSerLeuAspLeuAspPro 121  
DB 181 TGGATGAGCCACGCTGCTGGCCATACGAGATATCCAGATGACTCTCTGACCTAAACCA 240  
OY 122 SerMetLeuAspSerIleTrpLysProAspLeuPhePheAlaAsnGluLysGlyAlaAsn 141  
DB 241 TCCATGCTCGATCATCTGGAACCAACGCTCTTCTTCCCAATGAGAAAGGGCCAAAC 300  
OY 142 PheHisGluValThrThraspAsnLysLeuArgGlyIlePheLysAsnGlyAsnValLeu 161  
DB 301 TTCCATGAGGTGACCAACAGCAAGATTACGGCATCTTCAAGAATGGAAATGTCTC 360  
OY 162 TyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePromet 181  
DB 361 TACAGATAGATGATGACCTCATTTTGTCTGCCCAATGGACCTCAAGAACTTCCCAATG 420  
OY 182 AspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuPro 201  
DB 421 GACATACAGACCTGTACATGACAGCTGGAA----- 450  
OY 202 SerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTrpLeuGluAsp 221  
DB 451 -----AGCTTTGGCTACACCAATGATGACCTATGTTTAGTGCGCTAGAAAGT 498  
OY 222 AlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGlu 241  
DB 499 GCGCCGCTGTCGCAAGTGGCTGAGGGTTGACTGTACCCCAATTTATTTGGCGGATGAG 558  
OY 242 LysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluVal 261  
DB 559 AAGGATCTAGGGGATATGTACCAAGCACTCAATACAGGAAATTCACCTGCATTGAGGTA 618  
OY 262 LysPheHisLeuGlnLysGlnMetGlyTyrTyrLeuIleGlnMetTyrIlePheSerLeu 281

DB 619 AAGTTTACCCTGGAGCGGACAGATGGGCTACTGATCTGATCAATGATCATCCAGCCTA 678  
OY 282 LeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaArg 301  
DB 679 CTCATGCTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738  
OY 302 ValGlyLeuGlyIleThrThrValLeuThrIleThrThrGlnSerSerGlySerArgAla 321  
DB 739 GTAGGCTTGGCATCACACAGCTCTCCATGACCAATCAGACCTCTGGATCCAGGGCC 798  
OY 322 SerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeu 341  
DB 799 TCTTTGCTTAAGGTGCTCTAGGTAAAGCAATTGACATCTGATGCTGATGCTGCTGCTC 858  
OY 342 PheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPhelaSerArgGlnHisLys 361  
DB 859 TTTGTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 918  
OY 362 GluPheIleArgLeuArgArgArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArg 381  
DB 919 GAATTATGAGACTTCTGTAGAGGACGACGACGACGACGACGACGACGACGACGACGACG 978  
OY 382 GluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAspGly 401  
DB 979 GAGAGCCGCTTCTATTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1038  
OY 402 GlyProMetGluGlySerGlyIleTyrSerProGlnProAlaProLeuLeuArgGlu 421  
DB 1039 GGTCCAAATGGAAGGTTCAGACATTTATGATGCCCAACCCCAACCCCTTCTTAAGGAA 1098  
OY 422 GlyGluThrThrArgLysLeuTyrValAsp 431  
DB 1099 GGAGAACCATGAGGAACCTGATGTGGAC 1128

RESULT 4  
DRE404970 2045 bp mRNA linear VRT 21-NOV-2000  
LOCUS DRE404970  
DEFINITION Danio rerio mRNA for glycine receptor alpha2 subunit (glyr alpha  
LOCUS  
DEFINITION  
ACCESSION AJ404970  
VERSION AJ404970.1 GI:11322383  
KEYWORDS glycine receptor alpha2 subunit; glyr gene.  
SOURCE zebrafish.  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 2045)  
Imboden, M., de Saint Jan, D., Leulier, F., Korn, H., Goblet, C. and  
Bregestovski, P.  
Isolation and Characterization of a alpha 2-type zebrafish Glycine  
Receptor Subunit  
Unpublished  
2 (bases 1 to 2045)  
Goblet, C.  
Direct Submission  
Submitted 25 rue du Dr Roux, 75724-Paris Cedex 15, FRANCE  
Pasteur, Location/Qualifiers  
FEATURES  
source  
1. 2045  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/tissue-type="brain"  
/dev\_stage="adult"  
1. 2045  
/gene="glyr alpha"  
1. 519  
/gene="glyr alpha"  
520..1850  
/gene="glyr alpha"  
/codon\_start=1  
/product="glycine receptor alpha2 subunit"



```

/protein_id="CAC16687.1"
/db_xref="GI:11322384"
/translation="MTFSLITDNCISRCYHVFSGPESISLPLCFVLYMOKMSPS
DPLDKMGRTSGYDARIPRNFKGPYNNATNIFNSFGSTETTHMYRLNVRROOM
DPLATSEYRPSADLIDPSKMDITKPDLEFANEKANPHEVTDKLNLFONGVNL
YSRLTLIISCPMDLKNFPMIDQTCMQLSEFGYTNMDLFEVLSNPOVADLPLP
OFVLAKEKIDGVCYTKHYNTGKFCIEVKPHLEROMGYLYLOMYISPLLIVISWFEW
INMDAPARVGLGITVLTMTTOSGSRASLAPKYSYKALIDIMAVCLLIVFALLEY
AAVNFVSRHKEFIRLRKKORRORIEDIYRESRFPYFRGCGHOLKRDGDAVGS
SVAPPPYQVVLDEGAVKRRFYDRAKRIDITISRAVPLSLFNFYMTTYVNLNHE
DIRANP"
1891..2045
/gene="glyr alpha"
3'UTR
BASE COUNT      524 a      501 c      460 g      560 t
ORIGIN
Alignment Scores:
Pred. No.:      8,72e-168      Length:      2045
Score:          1749.50      Matches:      334
Percent Similarity: 84.91%      Conservative: 26
Best Local Similarity: 78.77%      Mismatches:  25
Query Match:    77.69%      Indels:      39
DB:             5      Gaps:      4
US-10-075-846-4 (1-431) x DREA04970 (1-2045)
Oy 9 LeuSerPheLeuLeuLeuTrpThrLeuProGlyGlnValLeuLeuArgValAlaLeuAla 28
Db 610 CTCTGCTTCTCTATCTCTGG----- 630
Oy 29 LysGluGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db 631 -----CAGAGCCATGTCCTCCATGGACTTTCG 660
Oy 49 AspLysLeuMetLysArgThrSerGlyTyrAspAlaArgLeuProAsnPhelyGly 68
Db 661 GACAACCTGATGGAGCAACCTCGGCTATGATGTCGATCAGACCACTTCAAGA 720
Oy 69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88
Db 721 CCACCGTAACCTCAGCTTATACATCTTATCAACAGTTTGGATCTATTACTGAGACA 760
Oy 89 ThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
Db 781 ACAATGGATACAGGCTAAACGCTTCTACGACAGAGAGGAGCAACCTAGACTGGCC 840
Oy 109 TyrArgGluTyrProAspSerSerLeuAspLeuAspProSerMetLeuAspSerIleTrp 128
Db 841 TACAGTGAATATCCCATGCACTCTAGACTGGACCTTCTATGTTGGACTCCATATGG 900
Oy 129 LysProAspLeuPhePheAlaAsnGluLysGlyValAlaAsnPhelyGlnValThrThrAsp 148
Db 901 AAACCTGACTGTTGTTTGGCTTAATGAGAAAGGTGCCAATGCAATGAGACCCACAGAC 960
Oy 149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeu 168
Db 961 AACACGCTGCTGAGATCTTTCAGAAATGGGAATGTGTTTACAGCAATCAGGTAACACTC 1020
Oy 169 IleLeuSerCysLeuMetAspLeuLysAsnPheproMetAspIleGlnThrCysThrMet 188
Db 1021 ATCCTCTCTGTCCTCCATGCACTGGAAGAAATTTCCCATGACATTCAGACTTACCATG 1080
Oy 189 GlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuValGly 208
Db 1081 CAGCTTGA-----ACGTTTGGC 1098
Oy 209 TyrThrMetLysAspLeuValPheGluTrpLeuGlnAspAlaProAlaValGlnValAla 228
Db 1099 TACACCATGAAACATGATCTTCGAGTGGCTTCTGATTAACCT---GTGCAAGTTGGC 1155
Oy 229 GluGlyLeuThrLeuProGlnIleLeuArgAspGluLysAspLeuGlyCysCysThr 248
Db 1156 GATGACTTGTCTCTCTCTGATTGACTAAAGAGAGAAAGATCTCGGCTACTGCACT 1215

```

```

Oy 249 LysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGln 268
Db 1216 AACCACTACAAACAGAGTAATTCACCTGCATTGAGGTGAAGTTTCACTTAGAGACAG 1275
Oy 269 MetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTrp 288
Db 1276 ATGGGCTACTATCTCATTCAGATGTAATCATCCAGCCTGCTGATTCATTTGTCCTGG 1335
Oy 289 ValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThr 308
Db 1336 GTGCTTTCGATCATATGATGATGCGGCACCGCTCGGCTGGGTTGGCTATCACACC 1395
Oy 309 ValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTrp 328
Db 1396 GTGCTGACATGACACCACTGAGCTGGTTCAAGACCTCGCTACCAAGTGTCTTAC 1455
Oy 329 ValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeu 348
Db 1456 GTGAGGCCATTTGACATCTGATGCGTGTTCGTCGTGATCGTTTGTGCTGATTTGG 1515
Oy 349 GluTyrAlaAlaIleAsnPhelySerArgGlnHisLysGluPheIleArgLeuArgArg 368
Db 1516 GACTATGCAACAGCACTTCGTCCTCAAGCAGACACAGAGAGTTCATCAGACTGCGCAAG 1575
Oy 369 ArgGlnArgArgGlnArgLeuGluGluAspIleIleGlnIleSerArg---PheTyrPhe 387
Db 1576 AACCAAGCAAGAGAGAGAAATGAGAGACCTTGTGAGAAAGCCGCTGCTTTTACTTC 1635
Oy 388 ArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArgAspLysGlyLysProMetGluGlySer 407
Db 1636 CGAGGGTACGAGCTGGGCGCATTCGCTTCCAAAACAAAGATGAGAGAGAGAGGGGCTCC 1695
Oy 408 GlyIleTyrSerProGlnProProAlaProLeuLeuArgGluGlyGluThrThrArgLys 427
Db 1696 AGGTATTCGCCCCACCACTCCAGTGCAGAGTCTTACAGCAGAGAGCGCTCGGAAG 1755
Oy 428 LeuTyrValAsp 431
Db 1756 CGCTTCTGTGAC 1767
RESULT 5
HSGLYRAL 1857 bp mRNA linear PRI 28-MAY-1993
LOCUS H sapiens alpha-2 strychnine binding subunit of inhibitory glycine
DEFINITION
ACCESSION X52008.1 GI:31848
VERSION glycine receptor; inhibitory glycine receptor; strychnine binding.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1857)
AUTHORS Greningloh,G., Schmieden,V., Schofield,P.R., Seeburg,P.H.,
Siddique,T., Mohandas,T.R., Becker,C.M. and Betz,H.
Alpha subunit variants of the human glycine receptor: primary
structures, functional expression and chromosomal localization of
the corresponding genes
EMBO J. 9 (3), 771-776 (1990)
JOURNAL MEDLINE 90183975
PUBMED 2155780
FEATURES
source
location/Qualifiers
1..1857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
/dev_stage="fetal"
388..1746
/gene="GLYRA1"
388..1746
/gene="GLYRA1"
/note="strychnine binding alpha-2 subunit"

```

```

/codon_start=1
/product="inhibitory glycine receptor"
/protein_id="CAA36257.1"
/db_xref="GI:31849"
/db_xref="SWISS-PROT:P23416"
/translation="MNRQLVNIILALPAFLETHNFRTRAECKDSDSRSKOPSTLSP
SDFLDKMGRSGTDARIRPNEKGPVNAVTCNIFINSGSQETETIMORVNIIFRQW
NDSRLAYSEYPPDSDLDLPSMLDSIMKPDLEFANEKANFEDVTDKLNIENKRY
LYSIRLTLLSCPMDLNFPMDVOTCTMQLSEFGYTMNDLFEWLSGPOVAEGLTL
POLIKKEKEGCTCKHNTGKFCIEVKFLEBROMGYLLIOMIPSLIYILSNVSP
WINDADPARVALGITVLTMTNOSGSRSLPKYSVKATIDIMAYCLLFFPALE
YAAVNFYSRQKEFLRLRRRKRONKEDYRESRNFPSGIMCHCLOVDKTRVKAT
PANLPQPKDGAIKRKFVDRARIDTISRAPLFLIFNIYITTKIRHEDVH
KK"
BASE COUNT      526 a      416 c      417 g      498 t
ORIGIN
Alignment Scores:
Pred. No.:      4,94e-161      Length:      1857
Score:          1682.50      Matches:      324
Percent Similarity: 86.10%      Conservative: 29
Best Local Similarity: 79.02%      Mismatches: 36
Query Match:    74.71%      Indels:      21
DB:             Gaps:      5
US-10-075-846-4 (1-431) x HSGLYRAL (1-1857)
OY 24 ArgValAlaLeuAlaValysGlu---GluValLysSerGlyThrLysGlySerGlnProMet 42
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 AGGACGGCTTCTCGCAAGACCATGACCTCCAGCTGTGAAACAACTTCACAGACCTTA 513
OY 43 SerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgIle 62
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 TCTCTTCAGATTTCTTGCAACAGTTAATGGGAAGACATCAGATATGATGCAAGAATC 573
OY 63 ArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPhe 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 574 AGGCCAAATTTAAAGGCTCCAGTAACCTTACTTCCCAATATTTTATCAACAGCTTTT 633
OY 83 SerSerValThrLysThrMetAspTyrArgValAsnValPheLeuArgGlnGlnIrrp 102
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 634 GGATCAGTCACAGAAAGACCATGGACGACGATATTTTCTGAGACAAACAGTGG 693
OY 103 AsnAspProAlaGluSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSer 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 694 AATGATTCACGCTGGCTGACGATGACGATGACGATGACCTCCGAGCTTGGACCATCC 753
OY 123 MetLeuAspSerIleTrrPlysProAspLeuPhePheIleAsnGlnLysGlyValAsnPhe 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 754 ATGCTAGACTCCATTGGAAACCAAGATTGTTCTTGGCAATGAGAAAGGTCCTCAACTTC 813
OY 143 HisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyr 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 814 CACGATGTCACACGACGACAAACAATTTGCTACGATTTGGAATAATGCAAGTGTCTAC 873
OY 163 SerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAsp 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 874 AGTATCAGACTCACCTTACCTTATCTCTCCATGACTTGAAGAAGAACTTTCCTGATGAT 933
OY 183 IleGlnIrrCysThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuProSer 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 934 GTCCAGACTGTACATCAGCTGAG----- 960
OY 203 LeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrrPLeuGluAspAla 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 961 -----AGTTTGGGTACACGATGATGACCTGATATTTGAGTGTGTTAAGTATGCT 1011
OY 223 ProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLys 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1012 CCA---GTGCAAGTTGCTGAAGATTGACCTGCCCACTTATTTTGAAGAAGAGAGAG 1068
OY 243 AspleuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLys 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 1069 GAACCTGGCTACTGTACAAACGCTACACACGTGAAAGTTTACCTGCATTTGAGCTAAG 11
OY 263 PheHisLeuGluArgGlnMetGlyTyrTyrThrLeuIleGlnMetCysTyrIleProSerLeu 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1129 TTTCACTCGGAAACGCCAAATGGATATTTATTTGATCCAGATGATACATCCCAAGCTCTGTT 1188
OY 283 IleValIleLeuSerTrrPValSerPheThrIleAsnMetAspAlaValProAlaArgVal 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1189 ATAGTAATTTTCTCTCGGTTCTCTTTTGGATTAATATGATGACAGCCCTGACAGGCTC 1248
OY 303 GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSer 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1249 GCACCTGGGATCACACAGTCTTAACGATGACCCAGACTTAAAGGCTCCAGGGCATCT 1308
OY 323 LeuProLysValSerTyrValLysAlaIleAspIleTrrPMetAlaValCysLeuLeuPhe 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1309 CTGCCAAAGTCTCTCTATGTAAGCAATTCATCTGATGATGCGGTGTCTGTGTT 1368
OY 343 ValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGlnHisLysGlu 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1369 GTGTTTGCTGCTTACTGCAATACGACGCGTGAACCTTCCTCCAGGCACACAGAGAG 1428
OY 363 PheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGlnArgIleIleGlnGlu 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1429 TTCTGGCCCTCCGAGAAAGACAGAGAGGCAATTAAGAAAGAGAGCTTACTCTGTGA 1488
OY 383 SerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArgAspGlyGly 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1489 AGTGTGTTTATTTATGAGCGTTATGAGATGGTGTGCTCCTCCCAAGTGAAGAGGAACA 1548
OY 403 ProMetGlnLysSerGlyIleTyrSer---ProGlnProProAlaProLeuLeuArgGlu 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1549 GCTGTCAGAGGCCACACCTGCCACACCCTCCACAGACCA-----AAAGAT 1596
OY 422 GlyGluThrThrArgLysLeuTyrValAsp 431
DB 1597 GGAGATGCTATCAAGAAAGATTTGTGGAC 1626

```

RESULT 6  
 RNO310837 1359 bp mRNA linear ROD 03-APR-2001  
 LOCUS Rattus norvegicus mRNA for glycine receptor alpha 2 precursor.  
 DEFINITION  
 ACCESSION AJ310837  
 VERSION AJ310837.1 GI:13548660  
 KEYWORDS glycine receptor alpha 2 precursor.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 1359)  
 Beato, M., Groot-Kormelink, P.J., Colquhoun, D. and Sivillotti, L.G.  
 Concentration dependence of single channel currents through rat  
 recombinant alpha 1 glycine receptors  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1359)  
 Groot-Kormelink, P.J.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-MAR-2001) Groot-Kormelink, P.J., Department of  
 Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,  
 London, WC1N 1AX, UNITED KINGDOM  
 FEATURES  
 source  
 1. 1359  
 /organism="Rattus norvegicus"  
 /strain="Sprague-Dawley males"  
 /db\_xref="taxon:10116"  
 /tissue\_type="whole brain"  
 /clone\_lib="rat brain marathon-ready cDNA"  
 /dev\_stage="8-12 weeks"  
 1. 1359  
 /function="ion channel"  
 /codon\_start=1  
 /evidence=experimental

```

/product="glycine receptor alpha 2 precursor"
/protein_id="CAC35981.1"
/db_xref="GI:13548661"
/translation="MNRQVNIILTALFAFLGTHNREARCKDHDSKHPQSOTLSP
SDFLDKMGRTSGDARIKPNFKGPPVAVNCNIFINSFSTVETETMDYRNIETLRQW
NDSRLAYSEYPPDSDLDPSMLDSIMKPDLEFANKEGANPHDVTDKLRIKNGKV
LYSLRLTLTSCPMDLKNPMDVQOTMOLESFGYTMNDLIFELISGPPQVAGTLF
POFLKEKEKELGCTKHNTGKFTCIKVFHLEROMGYLIIDIMYIPSLILVLSWSEF
MINMDAPARVAIGTITVLTMTQSSRASPISPVYKALIDIMAYCALIFEPALLE
YAAVNFYSROKHEFLRLRRQKRNKEEDTRESRENFSGCMGHCIQVKDGTAVKAT
PANLPQPPKDDADIKKKEVDRAKRIDITISNAEPLALFIFNIFWITVYKIRHEDVH
KK"
BASE COUNT      379 a      312 c      306 g      362 t
ORIGIN
Alignment Scores:
Pred. No.:      1.18e-160      Length:      1359
Score:          1677.00      Matches:      327
Percent Similarity: 82.91%      Conservative: 32
Best Local Similarity: 75.52%      Mismatches: 48
Query Match:    74.47%      Indels:      26
Db:             10      Gaps:      6
US-10-075-846-4 (1-431) x RNO310837 (1-1359)
Oy      1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlyGln 20
Db      13 CTACTGAAACATTTGGACAGCCTTGTTCGATTTTCTTA-----GGGACA 57
Oy      21 ValLeuLeuArgValAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySer 39
Db      58 AACCACTTCAAGGAGCAATTCGCAAGACCATGACGCTCGAAGAACATCCCTCG 117
Oy      40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsp 59
Db      118 CAGACCTGTCCTCTGATGATTTCTGTGATTAACATAATGGAGAACATCGATATGAT 177
Oy      60 AlaArgLLeaArgProAsnPhelLysGlyProProValAsnValThrCysAsnIlePheIle 79
Db      178 GCAGAGATCAGGCGCAATTTTAAAGTCCTCCAGTAAACGTTACTTGCATATTTTATTC 237
Oy      80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99
Db      238 AACAGTTTGGATGGTGCACAGAAACACCATGAGTACACGAGAGACATTTTCTGAGA 297
Oy      100 GlnGlnItrPasnAsnProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeu 119
Db      298 CAGCAGTGAACGATTCACGGCTGGCATACGATGATACCCAGATGATTCCTCGATTTG 357
Oy      120 AspProSerMetLeuAspSerIleTrpLysProAspLeuPhePheAlaAsnGluLysGly 139
Db      358 GATCCATCGATCGCGATTCGATTTGGAACACATTTGTTCTTTCGCAATGACAAAGGA 417
Oy      140 AlaAsnPhelLysGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159
Db      418 GCCAATTTTCATGATGTGCACACATGATTAACAAGTTGCTGGCGATTTCCAAATAAGGCAA 477
Oy      160 ValLeuLysSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhe 179
Db      478 GTGCTTACAGATTAATTAACCTACCTTATTCCTCCCAAGCAGCTGAAACACTTT 537
Oy      180 ProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerPro 199
Db      538 CCAATGATGATCGCAGACTGTCACAAATCAGCTGAG-----573
Oy      200 LeuProSerLeuSerLeuValGlyTyrThrMetLysAspLeuValPheGluTrpLeu 219
Db      574 -----AGTTTGGGTACACCATGATGATGACCTGATATTGAGTGGTTA 615
Oy      220 GluAspAlaProAlaValAlaGluGlyLeuThrLeuProGlnPheIleLeuArg 239
Db      616 AGTATGATGTCAC---GTACAAAGTTGTGAAGACTCACCCCTCGCTTATTTTGA 672

```

```

Oy      240 AspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIle 259
Db      673 GAAAGAGAGGAACTTGGCTATTCACAAAGCATTTACAACTGGCAAGTTTACCTGCATT 732
Oy      260 GluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIlePro 279
Db      733 GAGGTCAAGTTTACACCTGGAAGAGCGCAGATGGCGTATATTATTTGATCAGATGATATCC 792
Oy      280 SerLeuLeuIleValIleLeuSerTyrValSerPheTrpIleAsnMetAspAlaPro 299
Db      793 AGCGTGTGATAGTCATTTGCTCGCTCCCTTTTGATTAACATGAGTGCAGCTCCT 852
Oy      300 AlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrTrpGlnSerSerGlySer 319
Db      853 GCCAGGTTGGCTTGGCATTCACACCGCTCTGCAATGACTACCCAGAGTTCCAGTTTC 912
Oy      320 ArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValLys 339
Db      913 AGGCACTCTGCCCAAGGCTCATATGTGAACCAATTTGACATCTGGATGGACGTGTC 972
Oy      340 LeuLeuPheValPheAlaAlaLeuLeuGluTyrTrpAlaIleAsnPheValSerArgL 359
Db      973 CTCTGTTTGTGTTCCTGCTGCTTACTGGAATATCACACAGTGAACCTTGTCCAGGCAA 1032
Oy      360 HisLysGluPheIleArgLeuArgArgGlnArgTrpGlnArgLeuGluAspIle 379
Db      1033 CACAAAGAGTCTTCTGCTCCGAGACGACAGAGAGGCGAATAAGAAAGAGATGTT 1092
Oy      380 IleGlnLysSerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArg 399
Db      1093 ACTGCTGAAGACCGCTTTTAACTTACAGCGGTATGAGATGGGTCACTCCCTCAAGTGA 1152
Oy      400 AspGlyGlyProMetGluGlySerGlyIleTyrSer---ProGlnProProAlaProLeu 418
Db      1153 GATGTACAGCTGTCAAGGCTACACCTGCCAACCCACTTCCGCAACCCCA-----1203
Oy      419 LeuArgGluGlyGluThrThrArgLysLeuTyrValAsp 431
Db      1204 ---AAAGTGCAGATGCTATTCACAGACAGATTTGGAT 1239

```

**RESULT 7**  
**RNIGRA2**  
**LOCUS** RNIGRA2 3160 bp mRNA linear ROD 23-MAR-1993  
**DEFINITION** R. norvegicus mRNA for inhibitory glycine receptor alpha 2A subunit.  
**ACCESSION** X61159  
**VERSION** X61159.1 GI:288344  
**KEYWORDS** inhibitory glycine receptor alpha subunit.  
**SOURCE** Rattus norvegicus.  
**ORGANISM** Rattus norvegicus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
**REFERENCE** 1 (bases 1 to 3160)  
 Kunze, J., Kuryatov, A., Maule, Y., Malosio, M.L., Schmieden, V. and  
 Betz, H.  
 Alternative splicing generates two isoforms of the alpha 2 subunit  
 of the inhibitory glycine receptor  
**JOURNAL** FEBS Lett. 283 (1), 73-77 (1991)  
**MEDLINE** 91243883  
**PUBMED** 1645300  
**FEATURES**  
 source  
 1..3160  
 location/Qualifiers  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 552..1910  
 /codon\_start=1  
 /product="inhibitory glycine receptor alpha 2A subunit"  
 /protein\_id="CAA43471.1"  
 /db\_xref="GI:288345"  
 /db\_xref="SWISS-PROT:P22771"  
 /translation="MNRQVNIILTALFAFLGTHNREARCKDHDSKHPQSOTLSP  
 SDFLDKMGRTSGDARIKPNFKGPPVAVNCNIFINSFSTVETETMDYRNIETLRQW  
 NDSRLAYSEYPPDSDLDPSMLDSIMKPDLEFANKEGANPHDVTDKLRIKNGKV

LYSIRLRLTLSCPDMLNFMVDVOTCTMQLESFGYTMNDLIFENLSDGPVQVAGELTL  
 POFILKEKEELGYCTKHYNTGCTCIETKIEKFLEROMGYLLQMTYIPSLIYLISWSP  
 WINDADAPARVALGITTIVLTMTTOSSGRSALPKVSYKALDIIMAVCLLEFPVALLLE  
 YAAVNFVSRHKEFLRLRRRQKRONKEDYRESRPNQSGVGMCHLOVKRGTVKAT  
 PANLPQPPKADAIKKKFKVDRAKRIDTISNAEPFLFLININITYMTIKIRIEDVH  
 KK\*

sig\_peptide 552. .632  
 mat\_peptide 630. .1907

BASE COUNT 902 a 669 c 645 g 944 t  
 ORIGIN /product="inhibitory glycine receptor alpha 2A subunit"

## Alignment Scores:

Pred. No.: 3.65e-160 Length: 3160  
 Score: 1677.00 Matches: 327  
 Percent Similarity: 82.91% Conservative: 32  
 Best Local Similarity: 75.52% Mismatches: 48  
 Query Match: 74.47% Indels: 26  
 DB: 10 Gaps: 6

US-10-075-846-4 (1-431) x RNIGRAA2 (1-3160)

QY 1 MethThrLeuValProAlaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlyGln 20  
 Db 564 CTAGGAACTTTTGACAGCCTTGTTCGATTTTCTTA-----GGGACA 608  
 QY 21 ValLeuLeuArgValAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySer 39  
 Db 609 AACCACTTCAGGAGGACATTCGCAAAAGACATCCAGTCTGGAACATCCCTCC 668  
 QY 40 GlnProMetSerProSerAspPheLeuLysLysLeuMetGlyArgThrSerGlyTyrAsp 59  
 Db 669 CAGACCCGTCTCCTTCAGATTTCTTGATTAACATAAGGAGGACATCAGATATAT 728  
 QY 60 AlaArgGlyLeuArgProAsnPhelSerGlyProProValAsnValThrCysAsnIlePheIle 79  
 Db 729 GCAAGAACTCAGGCCAAATTTTAAAGCTCTCCAGTAAACGTTACTTGCATATTTTATC 788  
 QY 80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99  
 Db 789 AACACTTTTGATCGGTGCACAGAAACACCATGAGTACCGAGTACATTTTCTGAGA 848  
 QY 100 GlnGlnTrpAsnAspProArgLeuSerTyrArgGluTyrTrpProAspSerSerLeuAspLeu 119  
 Db 849 CAGCGGTGAGACGATTCAGCGCTGGCATACAGTACCCAGATGATTCCTGGATTTG 908  
 QY 120 AspProSerMetLeuAspSerIleTrpLysProAspLeuPheAlaAsnGluLysGly 139  
 Db 909 GATCATCGATGCTGATTCGATTCGTTGGAACACAGATTTGTTCTTGCCATGAGAAAGA 968  
 QY 140 AlaAsnPhelAsnGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159  
 Db 969 GCCAATTTTCATGATGATCACCACCATATTAACAAGTTGCGCGGATTTCCAAAAATGGCAAA 1028  
 QY 160 ValLeuTyrSerIleArgLeuThrLeuLeuLeuSerCysLeuMetAspLeuLysAsnPhel 179  
 Db 1029 GTGCTCTACAGATTAAGACTCACCTTGACTTATCTCTGCTCCATGAGACTGAGAACTTT 1088  
 QY 180 ProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSerPro 199  
 Db 1089 CCAATGATGTCACAGACCTGTCATATGACAGCTGGAG----- 1124  
 QY 200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLeu 219  
 Db 1125 -----AGTTTGGGTACACCATGATGATGATATATTGAGTGGTTA 1166  
 QY 220 GluAspAlaProAlaValGlnValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArg 239  
 Db 1167 AGTGATGGTCCA---GTACAAAGTGTGAGAGACTCACCTCCCTCAGATTATTTGAAA 1223  
 QY 240 AspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIle 259  
 Db 1224 GAAGAGAGAGAACTTGGCTATTGACAAAGCATTTACACACTGGCAAGTTTACCTGCATT 1283

QY 260 GluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetCysTrpIlePro 279  
 Db 1284 GAGGTCAAGTTTCAACCTGAGACGCCAGATGGGCTATATTGTGATTCAGATGATATATCC 1343  
 QY 280 SerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaIlePro 299  
 Db 1344 AGCCTGTGATGATGATTTTGTCTCGGTCTTCCTTTGGATTAACATGATGAGTGCCTCT 1403  
 QY 300 AlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrTrpGlnSerSerGlySer 319  
 Db 1404 GCCAGAGTTGGCCTTGGCATCACACCTCTGCACATGATACCTACCAAGATTCAGTTCC 1463  
 QY 320 ArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValCys 339  
 Db 1464 AGGCATCTGTGCCCAAGAGTCTCATATGTGAAGCAATTCATCTGATGGAGTGTGC 1523  
 QY 340 LeuLeuPheValPheAlaAlaLeuLeuGlnLysTrpAlaAlaIleAsnPhelValSerArgGln 359  
 Db 1524 CTCTGTTGTGTGGTGGCTGCTTACTGGAATATGCGACGAGTACTTGTCTCCAGGCCAA 1583  
 QY 360 HisLysGluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluGluAspIle 379  
 Db 1584 CACAAAGAGTTCCTTCCTCGTCCGGAGACAGACAGAAAGGCAAGATTAAGGAAGATGTT 1643  
 QY 380 IleGlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArg 399  
 Db 1644 ACTCTGTGAAGCCGTTTAACTTCAAGCGTATATGAGATGGGTCACTGCTCCAGTGAAA 1703  
 QY 400 AspGlyLysPrometGluGlySerGlyIleTyrSer---ProGlnProProAlaProLeu 418  
 Db 1704 GATGTTACAGCTGTGAAGCTTACACCTGCAACCCACTTCCGCAACCCCA----- 1754  
 QY 419 LeuArgGluGlyGluThrThrArgLysLeuTyrValAsp 431  
 Db 1755 ---AAGATGCAAGTGTATCAAGAAAGATTGTGAT 1790

RESULT 8  
 RNNEOGLY 3865 bp mRNA linear ROD 07-AUG-1991  
 LOCUS Rat NG1YR mRNA for neonatal glycine receptor.  
 DEFINITION X57281  
 ACCESSION X57281.1 GI:56743  
 VERSION 1  
 KEYWORDS glycine receptor; NG1YR gene.  
 SOURCE Rattus norvegicus.  
 ORGANISM Rattus norvegicus.  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 3865)  
 AUTHORS Hishinuma, F.  
 DIRECT SUBMISSION  
 JOURNAL Submitted (28-JAN-1991) F. Hishinuma, Mitsubishi Kasel Institute of Life Sciences, 11 Minamiooya, Machida-shi, Tokyo 194, Japan  
 REFERENCE 2 (bases 1 to 3865)  
 AUTHORS Akagi, H., Hirai, K. and Hishinuma, F.  
 TITLE Cloning of a glycine receptor subtype expressed in rat brain and spinal cord during a specific period of neuronal development  
 JOURNAL FEBS Lett. 281 (1-2), 160-166 (1991)  
 MEDLINE 91200276  
 PUBMED 1707830

FEATURES  
 source location/Qualifiers  
 1..3865 /organism="Rattus norvegicus"  
 /strain="Wistar"  
 /db\_xref="taxon:10116"  
 /tissue\_type="spinal cord"  
 /dev\_stage="10 days old"  
 1..3865 /gene="NG1YR"  
 1..3865 /gene="NG1YR"  
 /evidence="experimental"

CDS  
1173..2531  
/gene="NGlyR"  
/codon\_start=1  
/product="neonatal type of glycine receptor"  
/protein\_id="CAA0549.1"  
/db\_xref="GI:56744"  
/translation="NMROVLNLTALFAFLGNTNHFREAFCCKDHSRSGKHPQSLSP  
SDFDLKMGTSQDARIRPNFKGPNVTGNIINSFGSVETETMDYRNIELROOM  
NDSRLAYSEYPDSDLDPSMDSITKRPDLFANERKAGHNDITNKLIRISNKKV  
LYSIRRTLTSLSCPMDLAKNPMVDQCTMQLSESTYTMNDLIFEMLSGPRQVAGLTL  
PQTLKEEKLGTCTAHNTGKFTCLIEVFKHLEROMGYLIOMYIPSLILVILSWSE  
WIMMDAPARVALGITTYLMTQSSGSASRLSIDTMAVLCFLYFALVE  
YAAVNVSRQHEFLRLRRQKRONKEEDVRESNFGYGHCLQYVGDYAVKAT  
PAMPPLQPRDADAIKKKFFVDRAKRIDTISRAPLAFLENFYFWITTKIINHEVH  
KK"

BASE COUNT 1087 a 833 c 889 g 1056 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4,78e-160 Length: 3865  
Score: 1677.00 Matches: 327  
Percent Similarity: 82.91% Conservative: 32  
Best Local Similarity: 75.52% Mismatches: 48  
Query Match: 74.47% Indels: 26  
DB: 10 Gaps: 6

US-10-075-846-4 (1-431) x RNNEGLY (1-3865)

QY 1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuLeuThrPthrLeuProGlyGln 20  
Db 1185 CTAGTGACACTTTTGGACAGCCCTTGTGTTTCTTTCTTA-----GGGACA 1229

QY 21 ValLeuLeuArgValAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySer 39  
Db 1230 AACCACTTCAGGAGACATTCCTCAAGACCATGACTCCAGGCTCGAAGAACATCCCTCG 1289

QY 40 GlnPrometSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsp 59  
Db 1290 CACAGCCCTGCTCCTTCAGATTCTTGATTAACCTAATGGAGACATCAGCATATGAT 1349

QY 60 AlaArgLysArgProAsnPhelysGlyProProValAsnValThrCysAsnIlePheIle 79  
Db 1350 GCACAGATCAGGCCCAATTTTAAAGGCTCCCACTAAACCTTACTTCATATTATTATC 1409

QY 80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99  
Db 1410 AACAGTTTGGATCGGTACAGAAACCATGACTACCGAGTGAACATTTTCTGAGA 1469

QY 100 GlnGlnTyrPasnAspProArgLeuSerTyrArgLysTyrProAspAspSerLeuAspLeu 119  
Db 1470 CACAGTGGAGACATTCACGGCTGGCATACAGTACAGTACCATGATTCCTCGATTTG 1529

QY 120 AspProSerMetLeuAspSerIleTyrLysProAspLeuPheLeuAsnGluLysGly 139  
Db 1530 GATCCATCGATCGTGGATTGATTGAAACCATGATTGTTCTTCCCATGAGAAAGA 1589

QY 140 AlaAsnPheIleGluValThrThrAspAsnLysLeuLeuArgLysPheLysAsnGlyAsn 159  
Db 1590 GCCAATTTCCATGATGTCACCATGATACCAAGTGTGCGGATTGCCAAAATGCGAAA 1649

QY 160 ValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhe 179  
Db 1650 GTGCTGTACAGTATTTAGACTCACCTTGACTTATCCGTCCCATGACCTCGAAGAACTTT 1709

QY 180 PrometAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerPro 199  
Db 1710 CCATATGATTCAGACCTGTACAAATGCCACTGAG----- 1745

QY 200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrLeu 219  
Db 1746 -----AGTTTGGGTACACCATGATGACCTGATATTGAGTGTTA 1787

QY 220 GluAspAlaPheProAlaValGlnValAlaGlyLeuThrLeuProGlnPheIleLeuArg 239  
Db 1788 AGTGATGATGCCA---GTACAGATTGGCTGAAGAGCTACACCTGGCTGATTTATTGAAA 1844

QY 240 AspGluLysAspLeuGlyCysThrLysHisTyrAsnThrGlyLysPheThrCysIle 259  
Db 1845 GAAGAGACGACACTTGCTATTCACAAACCATATACACCTGCAAGTATTACCTGCATT 1904

QY 260 GluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuLeuIleGlnMetCylIlePro 279  
Db 1905 GAGGTACAGTTTACCTCGACACCCAGATGGCTATTATTGATCCAGATGATATCCCC 1964

QY 280 SerLeuLeuIleValIleLeuSerTyrPheSerPheThrPheLeuMetAspAlaIlePro 299  
Db 1965 AGCCTGCTGATCATCTATTGCTCTGCTCTCTCTTTGGATTAACATGATGACGCTCT 2024

QY 300 AlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerGlySer 319  
Db 2025 GCCAGGCTTCCCTTGGCATCACACCCCTCTGACATGACTACCCAGACTTCAGCTTCC 2084

QY 320 ArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCys 339  
Db 2085 AGGGCATCTTGCCTCAAGGCTCATATGTAAGCAATTCATGATGATGCGATGCGTGC 2144

QY 340 LeuLeuPheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGln 359  
Db 2145 CTTCCTGTTTGTGTTTGTGCTTACTGATATGACACATGAACTTGTCTCCAGGCAA 2204

QY 360 HisLysGluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluAspIle 379  
Db 2205 CACACAGAGATCTCTCTCTCCCGACAGACAGACAGGCGAATAAGCAAGAGATGTT 2264

QY 380 IleGlnIleSerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArg 399  
Db 2265 ACTCGTAAAGCCGTTTAACTTACCTACCGGTTATGGAGTGGCTACCTCCACAGTAAA 2324

QY 400 AspGlyLysPheMetGluGlySerGlyIleTyrSer---ProGlnProProAlaProLeu 418  
Db 2325 GATGTACACCTGCTCAAGGCTACACCTGCCACACCTTCCGCAACCCCA----- 2375

QY 419 LeuArgGluGlyLysThrThrArgLysLeuTyrValAsp 431  
Db 2376 ---AAGATCAGATGCTATCAGAGAGAAAGTGTGTGAT 2411

RESULT 9  
LOCUS HSU93917 3069 bp mRNA linear PRI 27-JUL-1998  
DEFINITION Human glycine receptor alpha 3 subunit mRNA, complete cds.  
ACCESSION U93917  
VERSION 093917.1 GI:3342235  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3069)  
AUTHORS Nikolic,Z., Laube,B., Weber,R.G., Licher,P., Kioschis,P.,  
Rouska,A., Mulhardt,C. and Becker,C.M.  
The human glycine receptor subunit alpha3. Glna3 gene structure,  
chromosomal localization, and functional characterization of  
alternative transcripts  
J. Biol. Chem. 273 (31), 19708-19714 (1998)  
MEDLINE 98344067  
PUBMED 9677400  
REFERENCE 2 (bases 1 to 3069)  
AUTHORS Nikolic,Z.  
TITLE Direct Submission  
SUBMITTER Submitted (17-MAR-1997) Biochemistry, University  
Erlangen-Nuremberg, Fahrstr. 17, Erlangen 91054, Germany  
LOCATION/Qualifiers  
1..3069  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

```

CDS
421..1770
/note="ligand gated ion channel"
/codon_start=1
/product="glycine receptor alpha 3 subunit"
/protein_id="AAC39917.1"
/db_xref="GI:3342236"
/translation="MAHRRHRLVSGFYWEAALLSLVATKETSARSRSAPMS
DFDLKMGRTSGDARIHPNFKGPVAVTCNIFINSFGSLAETMDYRVNIFLRKWN
DPLAYSEYDDSLDLPMSIMKPDLEFANEGANFHEVDNKLRFKGNVLY
YSIRLTLSCPMDLKNPMDVOTCIOMLESGFTYMDLFEOMDEAPVOYAEGLTLP
OLIKERKDLRYCTKHNTGKFCIEVRFBROMGYLLIOMYPSLIIYFSLVSVSF
INMAARARVALGTYLTMTTOSGSRASIPKSYKADIDMAVCLLYFSLVSEY
AANFVRKHRLERFRKRNRKNDDEVRESFSTAGMGPCLDAGKGMTPKGNHLY
QVMKPSDEMKKVFIDRAKKIDTISRACFLAFLIFNIYFWVIYKILRHEDHQOD"
BASE COUNT      962 a      547 c      583 g      977 t
ORIGIN
Alignment Scores:
Pred. No.:      2,86e-153      Length:      3069
Score:          1609.00      Matches:      310
Percent Similarity: 83.94%      Conservative: 35
Best Local Similarity: 75.43%      Mismatches:   42
Query Match:    71.45%      Indels:      24
DB:             9      Gaps:      5
US-10-075-846-4 (1-431) x HS093917 (1-3069)
Oy 22 LeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnPro 41
Db 481 CTGTACACCACTTTGGTTCACCAAGAAACAGACAGTGCAGATCTCGAAGTCTCCA 540
Oy 42 MetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArg 61
Db 541 ATGCACCTCTGATTTCTGATTAATTAATGAGCAGACATCATTAATTCATCAACAGT 600
Oy 62 IleArgProAsnPhelYsGlyProProValAsnValThrCysAsnIlePheIleAsnSer 81
Db 601 ATCAGACCCCAATTTAAAGCCCTCCAGTATATGTCACATCAACATATTCATCAACAGT 660
Oy 82 PheSerSerValThrLysThrMetAspTyrArgValAsnValPheLeuArgGlnGln 101
Db 661 TTGGGCTCTATCGAGACAGACACCATGATACAGAGATGATCTTTCTTCGTCAGAA 720
Oy 102 TrpAsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspPro 121
Db 721 TGGATGATCCCGCCCTCGTACAGTGAATATCTCTGACGACTTTAGACCTGACCCC 780
Oy 122 SerMetLeuAspSerIleTyrLysProAspLeuPhePheIleAsnGluLysGlyValAsn 141
Db 781 TCCATGTTGACCTCATTTGGAACCTGATTTGTTCTTGGCAATGAAGAGGTGCCAAC 840
Oy 142 PheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeu 161
Db 841 TTTTCATGAGTCACACAGACACAAATGCTAGAAATTTCAAAAAATGGAATGTTCTT 900
Oy 162 TyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMet 181
Db 901 TATTCATTAAGATTAACATTAACCTTCTCTGCAATGAGATTCACAAATTTTCCCATG 960
Oy 182 AspIleGlnThrCysThrMetGlnLeuGlnLysSerSerIleLeuCysSerProLeuPro 201
Db 961 GATGACAAACATGATATGCAACTGAA----- 990
Oy 202 SerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTyrLeuGlnLysP 221
Db 991 -----ACCTTGGGTACACATGATGATGATCTCATTTTGGATGGCAGATGAG 1038
Oy 222 AlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGlu 241
Db 1039 GCACCC---GTACAGAGTGGCAGAAAGACTCTTGGCCCGGTTCTGTTGAAGAAAGAA 1095
Oy 242 LysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluVal 261
Db 1096 AAGATTATACGATCTGCACTAAACATTACAAATACAGAAAGTTTACGTATAGAAGTG 1155
Oy 262 LysPheHisLeuGluArgGlnMetGlyTyrTyrThrLeuIleGlnMetTyrIleProSerLeu 281
Db 1156 CGATTCCATCTGGAGCGCAACAATGGATACTATCTGATCCAGATGATCATTTCCAGTCTC 1215
Oy 282 LeuIleValIleLeuSerTyrPvalSerPheThrIleAsnMetAspAlaIleProAlaArg 301
Db 1216 CTGATTTGTTATTTATCTCTGGTCTTCTCTGATCAACATGATGACGACCGGCCAGG 1275
Oy 302 ValGlyLeuGlyIleThrThrValLeuThrMetThrThrCysLeuSerSerGlySerArgAla 321
Db 1276 GTAGCTCTGGGATACCAACCGTCTAGCATGATGATGACAGATTCAGATTCAGAGCT 1335
Oy 322 SerLeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCysLeuLeu 341
Db 1336 TCCTTGGCAAAAGTTTCATATGTCAAAGCATATGATTTTGGATGGCAGATGCTCTT 1395
Oy 342 PheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGlnHisLys 361
Db 1396 TTGTGTTTTCAGCAGCTTCTGAGTATGACAGCTGTAATTTTGTATCAAGAACACAAA 1455
Oy 362 GluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluGluAspIleIleGln 381
Db 1456 GAACTTCTGAGATTTCAGACCAAGAGAAAGAAAT-----AAGGATGATGACGTAAAG 1506
Oy 382 GluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAspGly 401
Db 1507 GAAAGCCGATTCACCTTCACAGCCTTATGATGATGAGACCATGTCACAGCAAGAGATGCC 1566
Oy 402 GlyProMetGlyGlySerGlyIleTyrSerProGlnProProAlaPoleuLeu--Arg 420
Db 1567 ATGACTCCAAAGGAGC-----CCCAACCAACCTCTGTCAGTAATGCCAAA 1611
Oy 421 GluGlyGluThrThrArgLysLeuTyrValAsp 431
Db 1612 AGTCTGATGAATGAGAGAGGTCTTTATTCGAC 1644

```

```

Db 1096 AAGATTATACGATCTGCACTAAACATTACAAATACAGAAAGTTTACGTATAGAAGTG 1155
Oy 262 LysPheHisLeuGluArgGlnMetGlyTyrTyrThrLeuIleGlnMetTyrIleProSerLeu 281
Db 1156 CGATTCCATCTGGAGCGCAACAATGGATACTATCTGATCCAGATGATCATTTCCAGTCTC 1215
Oy 282 LeuIleValIleLeuSerTyrPvalSerPheThrIleAsnMetAspAlaIleProAlaArg 301
Db 1216 CTGATTTGTTATTTATCTCTGGTCTTCTCTGATCAACATGATGACGACCGGCCAGG 1275
Oy 302 ValGlyLeuGlyIleThrThrValLeuThrMetThrThrCysLeuSerSerGlySerArgAla 321
Db 1276 GTAGCTCTGGGATACCAACCGTCTAGCATGATGATGACAGATTCAGATTCAGAGCT 1335
Oy 322 SerLeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCysLeuLeu 341
Db 1336 TCCTTGGCAAAAGTTTCATATGTCAAAGCATATGATTTTGGATGGCAGATGCTCTT 1395
Oy 342 PheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArgGlnHisLys 361
Db 1396 TTGTGTTTTCAGCAGCTTCTGAGTATGACAGCTGTAATTTTGTATCAAGAACACAAA 1455
Oy 362 GluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluGluAspIleIleGln 381
Db 1456 GAACTTCTGAGATTTCAGACCAAGAGAAAGAAAT-----AAGGATGATGACGTAAAG 1506
Oy 382 GluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAspGly 401
Db 1507 GAAAGCCGATTCACCTTCACAGCCTTATGATGATGAGACCATGTCACAGCAAGAGATGCC 1566
Oy 402 GlyProMetGlyGlySerGlyIleTyrSerProGlnProProAlaPoleuLeu--Arg 420
Db 1567 ATGACTCCAAAGGAGC-----CCCAACCAACCTCTGTCAGTAATGCCAAA 1611
Oy 421 GluGlyGluThrThrArgLysLeuTyrValAsp 431
Db 1612 AGTCTGATGAATGAGAGAGGTCTTTATTCGAC 1644

RESULT 10
RNO310838
LOCUS      RNO310838      1443 bp      mRNA      linear      ROD 03-APR-2001
DEFINITION Rattus norvegicus mRNA for glycine receptor alpha.3 precursor.
ACCESSION  AJ310838
VERSION    AJ310838.1 GI:13548662
KEYWORDS   glycine receptor alpha 3 precursor.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 1443)
            Beato,M., Groot-kormelink,P.J., Colquhoun,D. and Sivajiotti,L.G.
            Concentration dependence of single channel currents through rat
            recombinant alpha 1 glycine receptors
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 1443)
            Groot-kormelink,P.J.
            Direct Submission
TITLE      Submitted (30-MAR-2001) Groot-kormelink P.J., Department of
            Pharmacology, The School of Pharmacy, 29/39, Brunswick Square,
            London, WC1N 1AX, UNITED KINGDOM
FEATURES   Location/Qualifiers
            1..1443
            /organism="Rattus norvegicus"
            /strain="Sprague-Dawley males"
            /db_xref="taxon:10116"
            /tissue_type="whole brain"
            /clone_lib="rat brain marathon-ready cDNA"
            /dev_stage="8-12 weeks"
            1..1443
            /function="ion channel"
            /codon_start=1
            /evidence=experimental

```

```

/product="glycine receptor alpha 3 precursor"
/protein_id="CAC35982.1"
/db_xref="GI:13548663"
/translation="MPWIRLFSSGSGISFRIMAHVHRHRTILSGFWEALLSLTAV
KENSARSRSRSPSPDLKLMKRTSGYARIRPNKGPVNTVCIFINSRSITAE
ITMDYRNIFLRQKMDPRLAYSEYDDSDLDLSDLSIMKPDLEFANKGANFHEV
ITDNKRLIRKNGNVLISIRLTLLSCPMKPNPMDQVCIMOLEFSGYIMDLIE
MODEAPQVAEGELTPQLLEKEDLCYKHYNGKTCIEVFEHLEROMGYLIDM
YIPSLIVLSWVSFIMMDAPARVALGTTVTMTTOSGSRVSLPKVSYKVAIDM
MAVCLIFVLSALLEVAANVAVNPSROHKEITVPRKRNKTEAPLEKFEPSDDDEY
RESRSFTAGCMKPCLOAKDGVYPRKPNHAYQVPRKSADEMRKYIDRAKKIDITISA
CFPLAFLEMFIFWVITKILRHEDIHQD"
BASE COUNT      405 a      320 c      324 g      394 t
ORIGIN

Alignment Scores:
Pred. No.:      1 37e-152      Length:      1443
Score:          1598.00      Matches:      314
Percent Similarity: 79.27%      Conservative: 34
Best Local Similarity: 71.53%      Mismatches: 49
Query Match:    70.96%      Indels:      42
DB:              10      Gaps:      7

US-10-075-846-4 (1-431) x RNO310838 (1-1443)

QY      9 LeuSerPheLeuLeuLeuTriPheLeuProGluValLeuLeuArgValAlaLeuAla 28
Db      79 CTTTCTGATTTTACTTCTGGAA-----GCCGACGCTTACTGAGATTGGTGCC 129
QY      29 LysGluLeuValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db      130 ACGAAGAAACAAACACGCGTCCCGAAGTCCCGCAAGTCACTTCGATTTCTG 189
QY      49 AspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLeuArgProAspPheLysGly 68
Db      190 GACAACTAATGGGAGAGACATCCGGTATGATCAGAAATCAGGCCCAACTTAAAGT 249
QY      69 ProProValasnValThrCysasnIlePheIleasnSerPheSerValThrLysThr 88
Db      250 CCTCCAGTTAATGTCACATGCAACATATTCTATAACAGCTTGGCTCCATCGCAGAGCG 309
QY      89 ThrMetAspTyrArgValasnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
Db      310 ACATATGATTACAGAGTAACATTTCTCTGTCAGAAAGTGAATGATCCTCGCCTTGA 369
QY      109 TyrArgGlnTyrProAspAspSerLeuAspLeuAspProSerMetIleAspSerIleTrp 128
Db      370 TACAGTGAATACCTGATGATTCATTAGACCTCGAACCCATTCATGTTGGACTCCATATG 429
QY      129 LysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGluValThrThrAsp 148
Db      430 AAACCTGACTTGTCTTGTCTAATAGAAAGGGGCTCACTCCACGAAGTCACACCGCAT 489
QY      149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeu 168
Db      490 AACCAAGTGTCTAAATAATTTCAAAATGGAATGCTTTTATTAATAAAGCTACATATTA 549
QY      169 IleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMet 188
Db      550 ACACTCTCTTGTCATATGATCTCAAGAAATTTCCCAATGATGTTCAAAACATGATATAG 609
QY      189 GlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208
Db      610 CAACTCGAA-----AGCTTTGGG 627
QY      209 TyrThrMetLysAspLeuValPheGlnTrpLeuGlnAspAlaProAlaValAlaGlnValAla 228
Db      628 TACACGATGATGATCTCATTTTGAATGCCAATGCAAGACCA--GTACAAGTGGCT 684
QY      229 GluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCysCysThr 248
Db      685 GAAGGACTACTTGGCTCTCAATTTCTGTTGAAGAAGAAAGATTGGCGACTGCACT 744

/product="glycine receptor alpha 3 precursor"
/protein_id="CAC35982.1"
/db_xref="GI:13548663"
/translation="MPWIRLFSSGSGISFRIMAHVHRHRTILSGFWEALLSLTAV
KENSARSRSRSPSPDLKLMKRTSGYARIRPNKGPVNTVCIFINSRSITAE
ITMDYRNIFLRQKMDPRLAYSEYDDSDLDLSDLSIMKPDLEFANKGANFHEV
ITDNKRLIRKNGNVLISIRLTLLSCPMKPNPMDQVCIMOLEFSGYIMDLIE
MODEAPQVAEGELTPQLLEKEDLCYKHYNGKTCIEVFEHLEROMGYLIDM
YIPSLIVLSWVSFIMMDAPARVALGTTVTMTTOSGSRVSLPKVSYKVAIDM
MAVCLIFVLSALLEVAANVAVNPSROHKEITVPRKRNKTEAPLEKFEPSDDDEY
RESRSFTAGCMKPCLOAKDGVYPRKPNHAYQVPRKSADEMRKYIDRAKKIDITISA
CFPLAFLEMFIFWVITKILRHEDIHQD"
BASE COUNT      405 a      320 c      324 g      394 t
ORIGIN

Alignment Scores:
Pred. No.:      1 37e-152      Length:      1443
Score:          1598.00      Matches:      314
Percent Similarity: 79.27%      Conservative: 34
Best Local Similarity: 71.53%      Mismatches: 49
Query Match:    70.96%      Indels:      42
DB:              10      Gaps:      7

US-10-075-846-4 (1-431) x RNO310838 (1-1443)

QY      9 LeuSerPheLeuLeuLeuTriPheLeuProGluValLeuLeuArgValAlaLeuAla 28
Db      79 CTTTCTGATTTTACTTCTGGAA-----GCCGACGCTTACTGAGATTGGTGCC 129
QY      29 LysGluLeuValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db      130 ACGAAGAAACAAACACGCGTCCCGAAGTCCCGCAAGTCACTTCGATTTCTG 189
QY      49 AspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLeuArgProAspPheLysGly 68
Db      190 GACAACTAATGGGAGAGACATCCGGTATGATCAGAAATCAGGCCCAACTTAAAGT 249
QY      69 ProProValasnValThrCysasnIlePheIleasnSerPheSerValThrLysThr 88
Db      250 CCTCCAGTTAATGTCACATGCAACATATTCTATAACAGCTTGGCTCCATCGCAGAGCG 309
QY      89 ThrMetAspTyrArgValasnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
Db      310 ACATATGATTACAGAGTAACATTTCTCTGTCAGAAAGTGAATGATCCTCGCCTTGA 369
QY      109 TyrArgGlnTyrProAspAspSerLeuAspLeuAspProSerMetIleAspSerIleTrp 128
Db      370 TACAGTGAATACCTGATGATTCATTAGACCTCGAACCCATTCATGTTGGACTCCATATG 429
QY      129 LysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGluValThrThrAsp 148
Db      430 AAACCTGACTTGTCTTGTCTAATAGAAAGGGGCTCACTCCACGAAGTCACACCGCAT 489
QY      149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeu 168
Db      490 AACCAAGTGTCTAAATAATTTCAAAATGGAATGCTTTTATTAATAAAGCTACATATTA 549
QY      169 IleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMet 188
Db      550 ACACTCTCTTGTCATATGATCTCAAGAAATTTCCCAATGATGTTCAAAACATGATATAG 609
QY      189 GlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208
Db      610 CAACTCGAA-----AGCTTTGGG 627
QY      209 TyrThrMetLysAspLeuValPheGlnTrpLeuGlnAspAlaProAlaValAlaGlnValAla 228
Db      628 TACACGATGATGATCTCATTTTGAATGCCAATGCAAGACCA--GTACAAGTGGCT 684
QY      229 GluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCysCysThr 248
Db      685 GAAGGACTACTTGGCTCTCAATTTCTGTTGAAGAAGAAAGATTGGCGACTGCACT 744

RESULT 11
LOCUS      RATIGRAS      2413 bp      mRNA      linear      ROD 07-MAR-1995
DEFINITION Rat inhibitory glycine receptor alpha subunit mRNA, complete cds.
ACCESSION M55250.1
VERSION M55250.1 GI:204882
KEYWORDS inhibitory glycine receptor.
SOURCE Rat (strain Wistar) adult brain, cDNA to mRNA, clone pGR48-alpha-3.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2413)
AUTHORS Kunze,J., Schmidtgen,V. and Betz,H.
TITLE Identification and functional expression of a novel ligand binding
subunit of the inhibitory glycine receptor
JOURNAL J. Biol. Chem. 265 (36), 22317-22320 (1990)
MEDLINE 91093073
PUBMED 2176214
FEATURES
Source
1..2413
location/Qualifiers
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/clone="pGR48-alpha-3"
/tissue_type="brain"
/dev_stage="adult"
1..2413
/gene="inhibitory glycine receptor"
400..1794
/gene="inhibitory glycine receptor"
/codon_start=1
/product="inhibitory glycine receptor alpha subunit"

```

```

/protein_id="AAAG3492.1"
/db_xref="GI:204883"
/translation="MAHVRERLTLSGFYMEALALSLVATKETNSARSRAPMS
DPFLDKLMRTSGYDARIRBNPKGPVAVTNNIFNSGSLAETMDVRVNILOKKN
DLRLAYSEYRDDSDLDPSMLDSIMKPDLEFANEGKAFHEVTTDNKLIRFKGNVL
YSTRLLTLSCPMDLKNFPMVOVTCIMQLSFGTYNDLLEFOWDEAPVOVABSLTL
QFLKEBKDLRYCTKHNKGTFCIEVRFHLEKMGYLLIOMYPSLLIYLSVSWF
INMDAPARVALGTTVLTMTQSSGSRAISLPKSYVKAIDIMAVCLLEFVSLLEY
AAVNFVSRKELERFRKRNKTEAFLEFYSPTDEVRSLSEFTFAYGMPCL
OAKGVVPGKGNHVVYMPKSADMRKVFIDRAKIDTISACPLAFLEIFNYIYWI
YKLRHEDIHQQD"
sig_peptide
400..498
/gene="inhibitory glycine receptor"
mat_peptide
496..1791
/gene="inhibitory glycine receptor"
/product="inhibitory glycine receptor alpha subunit"
BASE COUNT 697 a 503 c 557 g 656 t
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-151 Length: 2413
Score: 1592.00 Matches: 313
Percent Similarity: 79.04% Conservative: 34
Best Local Similarity: 71.30% Mismatches: 50
Query Match: 70.69% Indels: 42
DB: 10 Gaps: 7
US-10-075-846-4 (1-431) x RATIGRAS (1-2413)
OY 9 LeuserPheLeuLeuLeuTrpThrLeuProGlyGlnValLeuLeuArgValAlaLeuAla 28
Db 430 CTCTTGCGATTCTTCTGGGAA-----GCCGCCGCTGTACTAGTTGGTGCC 480
OY 29 LysGluGlnValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
Db 481 ACGAAGGAAACAAACAGTCAGCGTCCGAGAGTCCCAATGTCACCTTCGATTTTGTG 540
OY 49 AspLysLeuMetGlyArgThrSerGlyThrAspAlaArgLeuArgProAsnPhelLysGly 68
Db 541 GACAAACTAATGGGAGACATCCGGATGATGCAGAAATCAGGCCCACTTCAAAAGT 600
OY 69 ProProValAsnValThrCysAsnPhelPheLeuAsnSerPheSerValThrLysThr 88
Db 601 CCTCCAGTTAATGTCACATGCACATATTCATTAACACCTTGGCTCCATCGCAGAGAGC 660
OY 89 ThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
Db 661 ACTATGATTTACAGATAACATTTTCTCTGTCGAGAGTGAATGATCCTCGCTTGA 720
OY 109 TyrArgGlnTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerLettTrp 128
Db 721 TACAGTGAATACCCGATGATTCATTAAGCTCCGACCATGCTTGACTCCATATAGC 780
OY 129 LysProAspLeuPhePheAlaAsnGlnLysGlyAlaAsnPhelLysGlnValThrTrpAsp 148
Db 781 AAACCTGCTGTTGCTTGAATAGAGAGGGGCTTACCTCCAGAGATCCGCCGAT 840
OY 149 AsnLysLeuLeuArgGlyPheLysAsnGlnLysValLeuTyrSerLLeuArgLeuThrLeu 168
Db 841 AACCAAGCTGTAAGAAATTTTCAAAATATGAATGTTCTTTATTCATAAGGTGACATTA 900
OY 169 IleLeuSerCysLeuMetAspLeuLysAsnPhelProMetAspLLeuGlnThrCysThrMet 188
Db 901 ACACCTCTGCTGCCAATGATCTCAAGATTTCCCAATGATGTTCAAAACATGATATAG 960
OY 189 GlnLeuGlnSerSerLLeuLysSerProLeuProSerLeuSerLeuSerValGly 208
Db 961 CAACCTGCA-----AGCTTTGGG 978
OY 209 TyrThrMetLysAspLeuValPheGlnTyrPheGlnLysAspAlaProAlaValGlnValAla 228
Db 979 TACACGATGAATGATCTCATTTTGCATATGCGAAGATGAAGCACA---CTACAACTGCT 1035

```

```

OY 229 GlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLysAspLeuGlyCysCysThr 248
Db 1036 GAAGCACTCACTTGGCTTCATTTCTGTGAAAGAAAAAGATTTTGGCATCTGCACT 1095
OY 249 LysHisTyrAsnThrGlyLysPheThrCysLLeuValLysPheHisLeuGlnArgGln 268
Db 1096 AAACACTCAATATACAGAAAGATTACATGCTTAAAGATGATTCATCTTGACGGCNA 1155
OY 269 MetGlyTyrTLeuIleGlnMetTyrLLeuProSerLeuLeuIleValLLeuSerTrp 288
Db 1156 ATGGGCTACTCTGATGCCAATGATTCATTCGCCGCTTGATTCATCTGCTCTGG 1215
OY 289 ValSerPheTrpLLeuAsnMetAspAlaAlaProAlaArgValGlyLeuGlyLLeuThr 308
Db 1216 GTCTCATCTTCTGATTATACATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
OY 309 ValLeuThrMetTrpThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyr 328
Db 1276 GTACTTACGATGACCGCCAGAGTCTGATCCGGGCTTCTTTACCAAGGTCTCTAT 1335
OY 329 ValLysAlaIleAspLLeuTrpMetAlaValCysLeuLeuPheValPheAlaIleLeu 348
Db 1336 GTCAAGGCAATTTGACATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1395
OY 349 GluTyrAlaAlaIleAsnPhelValSerArgGlnHisLysGlnPheLLeuArgLeuArg 368
Db 1396 GAGTATGACGCGCGAATTTGTATCAAGCGCACAAAGAACTGATGAGTTTGGCGGA 1455
OY 369 ArgGlnArgArgGln-----ArgLeuGln----- 376
Db 1456 AAGAGGAAATTAACAGAAAGCTTTTGACATGAGAAAGTTTACCGTTTCTCAGACAG 1515
OY 377 GluAspLLeuIleGlnLysSerArgPheThrArgGlyTyrGlyLeuGlyHisCysLeu 396
Db 1516 GATGATGAGGAGGAGGAGGAGTCCGCTCAGCTTCACTGCTTATGAAATGAGGCCCTGCTT 1575
OY 397 GlnAlaArgAspGly-----GlyProMetGlnGlySerGlyLLeuTyrSerPro 412
Db 1576 CAAGCAAGAGATGAGTGTGTTCCAAAGAGTCCCAACCATGCTGTCCAGGTC----- 1626
OY 413 GlnProAlaProLeuLeuArgGlnGlyLLeuThrArgLysLeuTyrValAsp 431
Db 1627 -----ATGCCAAAGAGCGCCGATGAATAGAGMAGGTCTTATCATGCAC 1688
RESULT 12
HSGLYRA2 1715 bp mRNA linear PRI 28-MAY-1993
LOCUS H.sapiens alpha-1 strychnine binding subunit of inhibitory glycine
DEFINITION receptor mRNA.
ACCESSION X52009.1 GI:31850
VERSION X52009.1
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1715)
AUTHORS Greeningloh,G., Schimden,V., Schofield,P.R., Seeburg,P.H.,
Siddique,T., Mohandas,T.K., Becker,C.M. and Betz,H.
TITLE Alpha subunit variants of the human glycine receptor: primary
structures, functional expression and chromosomal localization of
the corresponding genes
JOURNAL EMBO J. 9 (3), 771-776 (1990)
MEDLINE 90183975
PUBMED 2155780
FEATURES
source
location/Qualifiers
1..1715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp21.2-p22.1"
/tissue_type="brain"
/dev_stage="fetal"

```





```
gene
1. 1744
/gene="Glr1"
229. 1602
/cds="Glr1"
/note="Glr1ins; alternatively spliced"
/codon_start=1
/product="glycine receptor alpha 1 subunit"
/protein_id="AAG14346.1"
/db_xref="GI:10180959"
/translation="MYSFNTLRVLYMNTIVFESLSAASKAEARSASKPMSPDELK
LMGRSGYDARIRPRFKGPPYVNSCNTINISFGSAETMTYRNIPIRQOMNDRLA
INERPDSDLIDPSKLDISIMKPDLEFANEKGAHHEITTDNKLKISNGVNVLSIRI
TLTLACPMDLNLFPMDOVTCIMOLESEFGYTNNDLIFEMOEGAAVADGLLPQILK
EEKDLRYCTKHNTGKFTFCIEARFHELRKMGVYLQMYIPSLIVLISMSIFWIMDA
APARVGLGITVLTSTOSGSRASLPVSYVKAIDIMAVCLLEFVSALLEVAYNV
VSRHKKELRRRRKRRHKSPLMNFQDEAGEGFENFSAYGMPCAOADGTSVKG
ANNSNTTPRPAKSPKSPKEMKLFQRAKKIDKISRIGFPAFLIFNMFTIYKIVR
REDVNO"
BASE COUNT      396 a      508 c      451 g      389 t
ORIGIN
Alignment Scores:
Pred. No.:      1,19e-150      Length:      1744
Score:          1580.00      Matches:      309
Percent Similarity: 80.89%      Conservative: 38
Best Local Similarity: 72.03%      Mismatches: 52
Query Match:    70.16%      Indels:      30
DB:             4      Gaps:      8
US-10-075-846-4 (1-431) x AF268375 (1-1744)
QY      12 LeuLeuLeuTPrThLeuProGlyGlnValLeuLeuArgValAlaLeuAlaGlyGln 31
DB      233 CTCCTACCTTTGGAG-----ACCATGTATCTTCTCAGTCTGTGCTTCCCAAG 303
QY      32 ValLysSerGlyThrLysGlySerGlnPrometSerProSerAspPheLeuAspLysLeu 51
DB      304 GCTGAAGCTGCTGCTGCTTCCCAAGCCCAAGTCACCGTCCGATTCTCGATTAAC 363
QY      52 MetGlyArgThrSerGlyTyrAspAlaArgIleArgProAspPheLysGlyProProVal 71
DB      364 ATGGGAGAGACTTCGTGATACGACGCCCAAGATCCCAATTTCAAGAGTCCCAATG 423
QY      72 AsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThrMetAsp 91
DB      424 AATGTCACCTGCACATTTTCATCAACAGCTTCGCTTCATTCGAGACCAACTATG 483
QY      92 TyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSerTyrArg 111
DB      484 TACAGGTCACATCTTCTCGGACGACGATGAGTACCCCGGCTGCTACATGAA 543
QY      112 TyrProAspSerSerLeuAspLysAspProSerMetLeuAspSerIleTyrLysProAsp 131
DB      544 TACCCCGATGACTCGCTGAGACTTGACCATTCATGCTGACTCATCTGGAACCTGAC 603
QY      132 LeuPhePheAlaAsnGlyGlyAlaAsnPheHisGlyValThrThAspAsnLysLeu 151
DB      604 CTGTTCTTGGCAATGAGAGAGGGGCCCATCTTCATGAGATCACACAGACCAACTG 663
QY      152 LeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSer 171
DB      664 CTGAGGATCTCCCGAAGCGGAATCTCTCTACACATTTAGATACCCCTGACATGGCC 723
QY      172 CysLeuMetAspLeuLysAsnPhePrometAspIleGlnThrCysThrMetGlnLeuGln 191
DB      724 TGCCCATGAGACTTGAAGAAATTTCCCATGATGCCAGAGGTGATATGAACTGAA 783
QY      192 SerSerSerIleLeuCysSerProLeuProSerLeuSerValGlyTyrThrMet 211
DB      784 -----AGCTTGGATACACCATG 801
QY      212 LysAspLeuValPheGlnTrpLeuGlnAspAlaProAlaValAlaGlnGlyLeu 231
```

```
DB      802 AACGACCTATCTTTGAGTGGCAGGACGAGGG---GCTGTGACAGTGGCAGATGACTA 858
QY      232 ThrLeuProGlnPheIleLeuArgAspGlnLysAspLeuGlyCysCysThrLysHisTyr 251
DB      859 ACTTACCCCAAGTTTATCTCTGAAAGAGAGAGAGACTTGAATATCTGCACCAACATTTAC 918
QY      252 AsnThrGlyLysPheThrCysIleGlnValLysPheHisLeuGlnIleArgGlnMetGlyTyr 271
DB      919 AACACAGCAATTCACCTGCATGAGAGCCCGGTTCCACCTGAGAGCGACAGATGGGCTAC 978
QY      272 TyrIleuIleGlnMetTyrIleProSerIleLeuIleValIleLeuSerTyrPhe 291
DB      979 TACCGATTCAGAGATGATCATCCAGCCTGCTCATTTGATCATCTCTGTGATCTCTTC 1038
QY      292 TrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrValLeuThr 311
DB      1039 TGGATCAACATGATGCTGCTCCGCCCGCTGGGCTGGGATCACCATGCTCTCAC 1098
QY      312 MetThrArgLysSerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAla 331
DB      1099 ATGACACCCAGAGCTCGGCTCGAGACATCCCGCCCAAGTATGATGAGAGGCC 1158
QY      332 IleAspIleTyrMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeuGlyTyrAla 351
DB      1159 ATCGACATTTGGATGGCCGCTGCTGCTCTTGTGTTCTCAGCCCTCTGTGAGTACGCC 1218
QY      352 AlaIleAsnPheValSerArgGlnHisLysGlnPheIleArgLeuArgArgArgGlnArg 371
DB      1219 GCCGTCATTTGTGTCTCTCGGACAGCAAAAGCTGCTCCATTCAGAGAGAGAGCAAGA 1278
QY      372 ArgGlnArg-----LeuGlnGlnAspIleIleGlnLysSerArgPhe 385
DB      1279 CATCACAGAGACCCCATGTAATCTGTCCAGAGAGATGAGCGTGGAGAGCGCCGCTTC 1338
QY      386 TyrPheArgGlyTyrGlyLeuGly---HisCysLeuGlnAlaArgAspGlyLysPromet 404
DB      1339 AACTTCTCGGCTTACGGATGGGCCACCTGCTTCCAGGCAAGAGGATCTGTCTC 1398
QY      405 GlnGly-----SerGlyIleTyrSerProGlnProProAlaProLeuLeuArgGln 421
DB      1399 AAGGCGCCCAACACAGACACACGACCAACCT---CCTCGGACCA---TCAAGTCC 1452
QY      422 GlyIleThrThrArgLysLeuTyrVal 430
DB      1453 CCGGAGAGATGCGAATAACTTCTATC 1479
RESULT 14
AF362764      1392 bp      mRNA      linear      ROD 09-AUG-2001
LOCUS      AF362764
DEFINITION      Mus musculus glycine receptor alpha 3 subunit (Glr3) mRNA,
complete cds.
ACCESSION      AF362764
VERSION      AF362764.1
KEYWORDS      GI:14010348
SOURCE      Mus musculus.
ORGANISM      Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 1392)
AUTHORS      Noegel,S., Becker,C. and Becker,K.
TITLE      Different glycine receptor isoforms are expressed in murine
cerebellum
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 1392)
AUTHORS      Noegel,S., Becker,C. and Becker,K.
TITLE      Direct Submission
JOURNAL      Submitted (21-MAR-2001) Biochemistry, University of Erlangen,
Fahrrstrasse 17, Erlangen 91054, Germany
FEATURES
source
1. 1392
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
```

```

/chromosome="8"
/map="26 cm"
/tissue_type="spinal cord"
1..1392
/gene="Glr3a"
1..1392
/gene="Glr3a"
/note="Ligand-gated ion channel"
/product="glycine receptor alpha 3 subunit"
/protein_id="AAK1962.1"
/db_xref="GI:14010349"
/translation="MAHVRHPRILVSGFYWEAALLSLVATKETSARSRAPMSGS
DFLDKLMKRTSGYDARIKPNFKGPVNVNINFINSGSIAETMDYRNVIPLRQKN
DPLRAYSYPDSDLDPSMDSIMKPDLEFANEKANHEVTDNKLIFKNGNVLV
SIRLTLLSCPMDKNFMDOYFCIMOLESEGYTMDLIFEMODEAPVOYAEGLTLPQ
FLKEEDLRVCYRKHYNTGKTCIEVFHLEBROMGYLIOMYIPLIIVLSVSMFI
NMDAAPARVALGTTDLTDTTQSSGSNASLPKSYVAIDIMAVCLLFYSALLETA
AVNFSRQHKELFRKRKRKNTKTEAFLEKFPYRSDDTDDVRESREFTAVGMPCLQ
AKGVVPRGNHAAVQVMPKSPDEKRVFIDRAKKIDTISRACEPLAFLINIFVWVIY
KILRHEDIHQOO"
BASE COUNT      388 a      310 c      313 g      381 t
ORIGIN
Alignment Scores:
Pred. No.:      1,25e-150      Length:      1392
Score:          1578.50      Matches:      317
Percent Similarity: 79.19%      Conservative: 33
Best Local Similarity: 71.72%      Mismatches: 50
Query Match:     70.09%      Indels:      42
DB:              10      Gaps:      9
US-10-075-846-4 (1-431) x AF362764 (1-1392)
OY      8 ThrleuSerPheleuLeuLeuTrp---ThrleuproglyGln-ValleuLeuArgValAl 26
Db      15 ACACCTTCGGACACTACTTCTTGATTTTACTTCTGGAGACCGCAGCTTGACTGACTTT 74
OY      26 AleuAlaLysGluGluValLysSerGlyThrLysGlySerGlnPrometSerProSerAs 46
Db      75 GGTCGCCACAAGGAACAAACAGTCGGCGATCTCGAAGTGCTCCATGTCACCTTCTCGA 134
OY      46 PheleuAspLysleuMetGlyArgThrSerGlyTyrAspAlaArgIleArgProAsnPh 66
Db      135 TTTTCTGCAAACTAATGCGGAGACATCGCGGTATGATGCAAGAATCAGACCACACTT 194
OY      66 elYsGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValTh 86
Db      195 CAAGGTCCTCCAGTATGTCACATGACATCATTAATTAACAGCTTTGGCTCCATTGC 254
OY      86 rlySThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProAr 106
Db      255 AGAGACGACATATGATATACAGAGTAACATTTTCTTCGTCAGAGAGTGAAATGATCTCG 314
OY      106 gLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSe 126
Db      315 TCTTGCACTAC---AGTATCTCTGACGATTCATTAAGACCTTGACCGCTATGTTGGATTC 371
OY      126 rIleTrpLysProAspLeuPhePheIleAsnGluLysGlyValAsnPhenHisGluValTh 146
Db      372 CATTGGAAAGCTGACTTGTCTTCTTAAGTAAGAGGGGCTTAACCTTCACGGAAGTCAC 431
OY      146 rThrAspAsnLysleuLeuValrGlePheLysAsnGlyAsnValLeuLysSerIleArgyle 166
Db      432 CACAGATTAACAAGCTGCTAAGATTTTCAAAAATGGAAGTCTCTTATTATCAATAAGGTT 491
OY      166 uThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhenProMetAspIleGlnThrCy 186
Db      492 GACATTAACACTCTCCCTCCAAAGATCTCAAAATAATTTCCCAATGATGATACAAACATG 551
OY      186 sThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeu 206
Db      552 CATTAATGCACACTTGA-----AG 569

```

```

OY      206 rValGlyTyrThrMetLysAspLeuValPheGluTrpLeuGluAspAlaProAlaValAl 226
Db      570 CTTTGGGTACACATGATGATGATCTCATTTTGAATGACGAGATGAAAGCCCA---GTACA 626
OY      226 nValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCy 246
Db      627 AGTGGCGAAGGACCTCTTGGCTCAATTTCTGTTTAAAGAAAGAAAGGATTCGCGTA 666
OY      246 sCysThrLysHisThrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGl 266
Db      687 CTGCACATAACACTACATATACAGGAAGTTTACTGATGATGAAGTGGATTCATCTTGA 746
OY      266 uArgGlnMetGlyTyrTyrLeuIleGlnMetCyrIleProSerLeuLeuIleValIleLe 286
Db      747 GCGTCATATGGGCGATTAATCACTGATCCAGATGATACATCCAGCCCTCTGATTCGTTATTC 806
OY      286 uSerTrpValSerPheThrPheLeuAsnMetAspAlaAlaProAlaArgValGlyLeuGly 306
Db      807 ATCTGGGCTCTCATTTCTGATTAACATGATGATGAGCCCGAGCTPAGACTTGGGCGAT 866
OY      306 eThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVa 326
Db      867 CACCACGATCTTACTATGACACACAGAGCTTGATCCCGGCGCTCTTACCAAGAGT 926
OY      326 lSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaAl 346
Db      927 GTGATACGTCACAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 986
OY      346 AleuLeuGluTyrAlaAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgyle 366
Db      987 ACTTCTGAGTATGACACCGCTGATTTTGTCTCAAGCACAACAGAGAGCTTTGAGCTT 1046
OY      366 uArgArgArgGlnArgArgGln-----ArgLeuGlu----- 376
Db      1047 TCGCGCAAGAGGAAGATTAAGACGGAAGCTTTTGCATGAGAGAGTTTACCCTTTCTC 1106
OY      377 -----GluAspIleIleGlnIleuSerArgPheTyrPheArgGlyTyrGlyLeuGlyH 394
Db      1107 AGACACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1166
OY      394 sCysLeuGlnAlaArgAspGly-----GlyPrometGluGlySerGlyIleTyr 410
Db      1167 CTGCTCTCAAGCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1224
OY      410 rSerProGlnProProAlaProLeuLeuArgGluGlyGluThrThrArgLysLeuTyrVa 430
Db      1225 -----ATGCCAAGAGCCCT-----GATGAATAGAGAAAGGCTTTCAT 1262
OY      430 lAsp 431
Db      1263 CGAC 1266

```

RESULT 15

DRE5812

LOCUS DRE5812 1628 bp mRNA linear VRT 01-AUG-1998

DEFINITION Danio rerio mRNA for glycine receptor (alpha1 subunit).

ACCESSION AJ005812

VERSION AJ005812.1 GI:3378595

KEYWORDS alpha1 subunit; glycine receptor.

SOURCE zebrafish.

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 1628)

AUTHORS David Watline,B.

TITLE Direct Submission

JOURNAL Submitted (27-APR-1998) David Watline B., Biotechnologies, Inserm U261 - Institut Pasteur, 25, Rue du Dr. Roux, 75724 Paris cedex15, FRANCE

REFERENCE 2 (bases 1 to 1628)

AUTHORS David-Watline,B., Goblet,C., de Saint Jan,D., Fucille,S.,

TITLE	JOURNAL FEATURES	Location/Qualifiers
Devignot, V., Bregestovski, P. and Korn, H. A novel alpha subunit of the glycine receptor from zebrafish.	primary structure and functional expression	Unpublished

Source

1. .1628

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 21:13:09 ; Search time 345.186 Seconds  
(without alignments)  
2811.853 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPAFLSLMTLMTLPCQ.....POPPAPLRCGETTRKLYVD 431

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL-frame+ p2n model -DEV-xip  
-O/cgnt2\_1/USPTO/spool/US10075646/runtat\_25062003\_163647-5141/app-query.fasta.1.782  
-DB-N.Geneseq.101002 -QFMT-fastlap -SUFFIX-p2n.rng -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR-SCORE-pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE-LOCAL -OUTFMT-rlc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075646.ecgn.1.1.511.arnat.25062003.163647-5141 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N.Geneseq.101002:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2163	96.0	1874	24 AAD33667	Human TRICH-22 CDN
2	2157	95.8	1600	21 AAD31678	CDNA sequence enco
3	984.5	43.7	657	13 AAO25722	GABA-A receptor be
4	871	38.7	2404	24 AB199254	Mouse ischaemic co
5	761	33.8	39796	21 AOC61681	Nucleotide sequenc
6	754.5	33.5	2400	22 AAD21397	R. sanguineus glut
7	750	33.3	4621	24 AAD26939	Heliothis virescen
8	749	33.3	1609	24 AAD26937	Heliothis sp. HEG3
9	742.5	33.0	1197	22 AAD21378	R. sanguineus LGIC
10	742.5	33.0	1197	22 AAD22082	R. sanguineus Gluc
11	742.5	33.0	2138	22 AAD21395	R. sanguineus glut
12	742.5	33.0	2289	22 AAD21396	R. sanguineus glut
13	737.5	32.7	1614	24 AAD22072	Dermacentor variab
14	737	32.7	1614	24 AAD22070	Dermacentor variab
15	732.5	32.5	1614	24 AAD22071	Dermacentor variab
16	731.5	32.5	1368	20 AAX24372	Cat flea glutamate
17	731.5	32.5	3958	17 AAT43596	Glutamate-gated ch
18	729	32.4	2340	23 AB116609	Drosophila melanog
19	728.5	32.3	1640	23 AAD26938	Heliothis sp. HEGM
20	728.5	32.3	1844	20 AAV64372	GABA-gated chlorid
21	726	32.2	2066	21 AAZ44048	D. simulans GABA r
22	725.5	32.2	1766	22 AAD04276	Short form of S. a
23	725.5	32.2	1766	22 AAD04277	Short form of S. a
24	724.5	32.2	1519	20 AAV64373	GABA-gated chlorid
25	724	32.1	1657	19 AAV55007	GABA-gated chlorid
26	724	32.1	1657	19 AAV55008	GABA-gated chlorid
27	720.5	32.0	3619	23 AB113035	Drosophila melanog
28	720	32.0	2066	21 AAZ44045	D. melanogaster po
29	720	32.0	2066	21 AAZ44046	D. melanogaster po
30	719	31.9	3442	22 AAD21373	Dermacentor variab
31	719	31.9	3598	22 AAD21372	Dermacentor variab
32	718	31.9	2066	21 AAZ44047	D. melanogaster GA
33	717.5	31.9	1823	19 AAX16065	DNA sequence of lu
34	717	31.8	1491	17 AAO99330	GABA receptor subu
35	717	31.8	1491	20 AAV84665	Insect GABA recept
36	715	31.7	1498	23 AB118879	Drosophila melanog
37	712	31.6	1866	14 AAO63131	GABA-A receptor be
38	712	31.6	1866	14 AAO63131	Human GABA recept
39	698.5	31.0	1640	22 AAI60040	Human polynucleoti
40	698.5	31.0	1594	22 AAI58254	Human polynucleoti
41	687.5	31.0	1594	24 AB199308	Mouse ischaemic co
42	694.5	30.8	1866	24 ABL67062	Thyroid cancer rel
43	690	30.6	1879	22 AAD04274	Long form of S. am
44	690	30.6	1879	22 AAD04275	Long form of S. am
45	687.5	30.5	1970	21 AAZ44049	D. melanogaster po

## ALIGNMENTS

RESULT 1  
AAD33667  
AAD33667 standard; CDNA: 1874 bp.  
AC  
XX  
AC  
XX  
DT  
XX  
DE  
XX  
XX  
Human TRICH-22 CDNA.  
Human: transporter and ion channel; TRICH-22; transport disorder; angina;  
amyotrophic lateral sclerosis; cystic fibrosis; neuromuscular disorder;  
cardiac disorder; polyomyelitis; diabetes; neurological disorder; cancer;  
depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;  
cell proliferated disorder; infertility; arteriosclerosis; gene therapy;  
Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;  
myasthenia gravis; multiple sclerosis; metabolic disorder; hypertension;  
acquired immune deficiency syndrome; immunological disorder; scleroderma;



```
Db 964 GAGAGAGCTAGAGCTGTTGTAACCAAGCACTACACACAGGGAATTCTGATCGAG 1023
Qy 261 VALLYSPhHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSer 280
Db 1024 GTAAGCTTCCACCTGGAAGGAGATGGGCTACTATCTATCTAGATGATGATCCAC 1083
Qy 281 LeuLeuIleValIleLeuSerTyrValSerPheTyrIleAsnMetAspAlaAlaProAla 300
Db 1084 CTACTCATGTCATCTGCTGCTGGGCTCTCTCTGATCAACATGATGATCTGCTGCTGCC 1143
Qy 301 ArgValGlyLeuGlyIleThrThrValIleuThrMetThrThrGlnSerSerGlySerArg 320
Db 1144 CCGTGCGGCTGGGCTGACCAACCGCTGCTACCATGACACCCAGAGCTTGCTGCCG 1203
Qy 321 AlaserLeuProLysValSerTyrValLysAlaIleAspIleTyrMetAlaValCysLeu 340
Db 1204 GCCTCTTTCCTTAGGTGTCCTACGTAAGGCAATGATGATGCTGCTGCTGCTG 1263
Qy 341 LeuPheValPheAlaAlaLeuLeuGluTyrTyrAlaAlaIleAsnPheValSerArgGlnHis 360
Db 1264 CTTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1323
Qy 361 LysGluPheIleArgLeuArgArgArgArgArgArgArgArgArgArgArgArgArg 380
Db 1324 AAGAATTCATACGACTTCGAAGAGGCAAGGCGCAACGCTTGAGAGAGATATCATC 1383
Qy 381 GlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAsp 400
Db 1384 CAAGAAAGCTGTTTCATTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1443
Qy 401 GlyGlyPheMetCysLeuGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg 420
Db 1444 GGAGGTCATAGGAGAGGTTCTGGCATTTATAGCCCAACCTCCAGCCCTCTTAAAG 1503
Qy 421 GlnGlyGluThrThrArgLysLeuTyrValAsp 431
Db 1504 GAGGAGAAACCGCGGAAACTCTACGTGAC 1536

RESULT 2
AAC61678
ID AAC61678 standard; DNA; 1600 BP.
AC AAC61678;
XX
XX 19-FEB-2001 (first entry)
DT
XX
XX cDNA sequence encoding a human ataxia protein.
DE
XX
XX Human; ataxia; gene therapy; ss.
KW
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1..1254
FT CDS /tag= a
FT /product= "ataxia protein"
XX
XX MO200058461-A1.
XX
XX PD 05-OCT-2000.
XX
XX PF 23-MAR-2000; 2000MO-EP02600.
XX
XX PR 26-MAR-1999; 99EP-0106343.
XX
XX PA (RAPF/) RAPOLD-HOERBRAND G.
XX
XX PI RAPOLD-Hoerbrand G;
XX
XX DR MPI; 2000-656166/63.
XX
XX DR P-PSDB; AAB19336.
XX
XX Novel nucleic acid sequence encoding human ataxia protein for screening
```

```
PT compounds useful for treating disorders relating to mutations in ataxia
PT gene
XX
XX Claim 2: Page 18-20; 47pp; English.
PS
XX
XX The present sequence encodes a human ataxia protein. The ataxia
CC protein and polynucleotides are useful for diagnosing and treating
CC disorders related to ataxia. Ataxia gene sequences are useful in
CC gene therapy, and as diagnostic tools or reagents for identifying and
CC characterizing genetic defect involved in the disorders and diseases
CC related to ataxia.
XX
XX Sequence 1600 BP; 378 A; 440 C; 379 G; 403 T; 0 other:
SQ
XX
XX Alignment Scores:
Pred. No.: 4,06e-240 Length: 1600
Score: 2157.00 Matches: 414
Percent Similarity: 96.29% Conservative: 1
Best Local Similarity: 96.06% Mismatches: 2
Query Match: 95.78% Indels: 14
DB: 21 Gaps: 1
US-10-075-846-4 (1-431) x AAC61678 (1-1600)
Qy 1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuTyrThrLeuProGlyGln 20
Db 1 ARGACAACTCTGCTGCTGCAACCCCTCTCTCTCTCTGAGACCTGCGAGGCGAG 60
Qy 21 ValLeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGln 40
Db 61 GTCCCTCCTCAGGCTGCGCTTGCGCAAAAGAGAACTGGAATCTGGAACCAAGGCTCCAG 120
Qy 41 PrometSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAla 60
Db 121 CCCATCTCCCTCTCTATTTCTCTAGACAACTTATGGGCGAATCTGATATGATGCC 180
Qy 61 ArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsn 80
Db 181 AGGATTCGGGCCCAATTTTAAAGGCCCAACCGTGAACGTAACCTTCATCATCAAC 240
Qy 81 SerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGln 100
Db 241 AGTTTCAGCTCCATCCACCAAGACCAATGATGATCGGCTGATGATCTTCTTGGCGGCA 300
Qy 101 GlnTyrAsnAspProArgLeuSerTyrArgGlyTyrProAspAspSerLeuAspLeuAsp 120
Db 301 CAGTGAATGACCCACGCTGCTCTACCGAATATCTGATGACTCTTGACCTCGAT 360
Qy 121 ProSerMetLeuAspSerIleTyrLysProAspLeuPheAlaAsnGluLysGlyAla 140
Db 361 CCTCCATGCTGACCTCTATCTGGAAGCCAGACCTCTCTCTGCTAATGAAAGGGGCC 420
Qy 141 AsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnVal 160
Db 421 AACTTCATGAGGTGACCAAGCAAGATTAAGTCACTTCGATCAAGATGGAATGTG 480
Qy 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePro 180
Db 481 CTGTACAGCATCAGGCTGACCTCATTTTGTCTGCTGCTGATGACCTCAAGAACTTCCCC 540
Qy 181 MetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuGlySerProLeu 200
Db 541 ATGACATCCACAGCTGACGATGCAAGCTTGAC----- 573
Qy 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluThrLeuGlu 220
Db 574 -----AGTTTGGCTACACCATGAAAGACCTCTGTTGAGTGGCTGAA 618
Qy 221 AspAlaProAlaValAlaGlnGluGlyLeuThrLeuProGlnPheIleLeuArgAsp 240
Db 619 GATGCTCTGCTCTCCAAAGTGGCTGAAGGGGCTGACTCTGCCCAAGTTATCTTGGCGGAT 678
Qy 241 GluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu 260
```

Db	679	GAGAAGGATCATGAGCTGTTGTTACCAAGACTACAAACAGGAAATTCACCTGCATCGAG	738
Qy	261	VallysPhehstleuGlutArgGlnmeGlytyrTyrIleuIleGlnmetTyrIleProSer	280
Db	739	GTAAGATTTCACCTCGGAACGGCAATGGGCTACATCATGTGATTCAGATGTACATCCCAAGC	798
Qy	281	LeuLeuIleValIleIleSerTyrValSerPheThrPileasmetAspAlaIleProIle	300
Db	799	CTACTCATCGTCATCCCTGTCTCTGGGTCTCCCTTCGTGATCAACATGATGATGCCCTGCC	858
Qy	301	ArgValGlyleuGlyIleThrValIleuThrmetThrThrGlnSerSerGlySerArg	320
Db	859	CGTGTGGCCCGGGGCATCACCCCGCTGCTCACCATGACCAACCAGAGCTCTGGCTCCCGG	918
Qy	321	AlaSerIleuPhePolsValSerTyrValIleValIleAspIleTyrMetAlaValCysIleu	340
Db	919	GCCTCTTGGCTTAAGGCTGCTCTACGTTGAAGCAATCGCATCTGATGGCTGTGTGTG	978
Qy	341	LeuPheValPheAlaIleIleleuGluTyrAlaIleAspIleAsnPheValSerArgGlnHis	360
Db	979	CTCTTGTGTGCTCGCTGCTCTGAGATGCTGCATTAATTTTGTTCGTGCACGAT	1038
Qy	361	LysGluPheIleArgLeuArgArgArgGlnArgGlnArgGlnArgLeuGluAspIleIle	380
Db	1039	AAAGAAATTCATACGACTTCGGAAGAAGCGAGCGCCCAACGCTTGAGAGAGATATCATTC	1098
Qy	381	GlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaArgAsp	400
Db	1099	CAAGAAGTCTTTTCTATTTCCGTGGCATGTGGCTTGGGCCACTGCTCTGCAGCAAGAT	1158
Qy	401	GlyGlyPheMetGluGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg	420
Db	1159	GGAGGTCCAAAGGAAGGTTTCGGCATTTATGTGTCGCCCAACCTCCAGCCCTCTTAAGG	1218
Qy	421	GluGlyGluThrThrArgGlyLeuTyrValAsp	431
Db	1219	GAAAGAGAAACACGCGGAACCTACTGCGAAC	1251
RESULT 3			
AAQ25722			
AAQ25722 standard; DNA; 657 BP.			
AAQ25722;			
AAQ25722;			
08-DEC-1992 (first entry)			
GABA-A receptor beta-subunit.			
Expression plasmid; gamma-aminobutyric acid; OmP; glycine; ss.			
Synthetic.			
JP04144683-A.			
19-MAY-1992.			
05-OCT-1990; 90JP-0267743.			
05-OCT-1990; 90JP-0267743.			
(MITU ) MITSUBISHI KASEI CORP.			
WPI; 1992-214122/26.			
Prepn. of N-terminal extracellular site protein - by culturing			
E.coli transformed by a plasmid comprising the tac promoter, ribosome			
binding site, etc.			
Claim 1; Page 2; 12pp; Japanese.			
The sequence given is the gamma-aminobutyric acid (GABA) A receptor			
beta-subunit. This sequence is used in an expression plasmid operably			

CC	linked to the tac promoter, a ribosome binding sequence, the E. coli
CC	outer membrane protein ompF signal peptide coding sequence and a
CC	sequence coding for the N-terminal extracellular site of either the
CC	gamma-aminobutyric acid (GABA) A receptor alpha subunit or the glycine
CC	receptor alpha subunit. (GABA A receptor alpha subunit can be used to
CC	transform E. coli to produce an N-terminal extracellular site protein
CC	of ion channel direct binding type receptor.
CC	
xx	
50	Sequence 657 BP: 171 A; 185 C; 156 G; 145 T; 0 other:
	Alignment Scores:
	Pred. NO.: 1.86e-104 Length: 657
	Score: 984.50 Matches: 182
	Percent Similarity: 87.34% Conservative: 18
	Best Local Similarity: 79.48% Mismatches: 14
	Query Match: 43.72% Indels: 15
	DB: 13 Gaps: 2
US-10-075-846-4 (1-431) x AAQ25722 (1-657)	
QY	40 GluPROMETserProSeraspPheLeuaspLysleuMetGlyArgThrSerGlyTyrASP 59
DB	16 AAGCTTATGTCACCCCTCGGACCTTCCGATAAGCTTATGGAGAGACTCTGGGTATGAT 75
QY	60 AlaArglleArgProaspNhelysGlyProProValasnaValThrCysAsnIlePheIle 79
DB	76 GCCGAGATCGAGACCACCTTAAAGTCTCCTCGTAACGTAAGTGCACATCTTCATC 135
QY	80 AsnSerPheSerSerValThrLysThrTrpMetAspTyrArgValAsnValPheLeuArg 99
DB	136 AACGCTTGTCGTTCTATGCCCGAGAACACCATGAGCTACAGGGTCAACATCTTCTAGG 195
QY	100 GluGlnTrpAsnAspProArgSerTyrArgGluTyrProaspPaspSerLeuaspLeu 119
DB	196 CAGCAGTGAAGAACGCCCGCTCGCGCTACATGATATACCTGACGACTCTCGACCTT 255
QY	120 AspProSerMetLeuaspSerIleTrpLysProaspLeuPhePheAlaAsnGluLysGly 139
DB	256 GACCACATCCATGTTGGATTCCATCTGGAAGCCCTGACTGTTCTTTGCCAATGACAAAGGG 315
QY	140 AlaAsnPheIstGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159
DB	316 GCCCACTTCCAGAGATTCACACGAGCAACACAGACGCTGAGAGATCTCCCGAGACGGCAC 375
QY	160 ValLeuTyrSerIleArgLeuThrIleuLeuSerCysLeuMetaspLeuLysAsnPhe 179
DB	376 GTCTCTACAGCATCGAATCACCCCTGAGCTCGGCTCCCATGAGACCTTCAACAAATTTT 435
QY	180 PheMetaspIleGlnThrCysThrMetGlnLeuGlnUserSerSerIleLeuCysSerPro 199
DB	436 CCGATGACGTACAGACATGTATCATGCACTGGAA----- 471
QY	200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysaspLeuValPheGluTrpLeu 219
DB	472 -----AGCTTTGGTATTACCATGAGACGACCTCATCTTTGAGTCG--- 510
QY	220 GluAspAlaIleProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArg 239
DB	511 GAAGAGCAAGACCTTGTGAGGGGAGATGGAGACGACCTCGCTCACTTATATCTTAAG 570
QY	240 AspGluLysaspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIle 259
DB	571 GAGSAGAGAGATCTGAGATGATCTGCACCAAGCACTACACACAGTAATTACCTGCATT 638
QY	260 GluValLysPheHisLeuGluArgGln 268
DB	631 GAGGCCGATTCACACCTGAGACGGAG 657
RESULT 4	
ABI99254	
ID	ABI99254 standard; cDNA; 2404 BP.
xx	
ABI99254;	





```

XX Nucleotide sequence of the human ataxia gene.
XX Human; ataxia; gene therapy; ss.
XX Homo sapiens.
XX Location/Qualifiers
XX Key 29850..29921
XX intron /tag= a
XX intron /tag= b
XX exon /tag= c
XX intron /tag= d
XX exon /tag= e
XX intron /tag= f
XX exon /tag= g
XX intron /tag= h
XX exon /tag= i
XX intron /tag= j
XX exon /tag= k
XX intron /tag= l
XX exon /tag= m
XX WO200058461-A1.
XX 05-OCT-2000.
XX 23-MAR-2000; 2000MO-EP02600.
XX 26-MAR-1999; 99EP-0106343.
XX (RAPP/) RAPPOLD-HOERBRAND G.
XX Rappold-Hoerbrand G;
XX MPI: 2000-656166/63.
XX Novel nucleic acid sequence encoding human ataxia protein for screening
XX compounds useful for treating disorders relating to mutations in ataxia
XX gene
XX
XX Claim 6, Page 22-44; 47pp; English.
XX
XX The present sequence represents the human ataxia gene. The ataxia
XX protein and polynucleotides are useful for diagnosing and treating
XX disorders related to ataxia. Ataxia gene sequences are useful in
XX gene therapy, and as diagnostic tools or reagents for identifying and
XX characterizing genetic defect involved in the disorders and diseases
XX related to ataxia.
XX
XX Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 other:
XX
Alignment Scores:
Pred. No.: 1,036-75 Length: 39796
Score: 761.00 Matches: 230
Percent Similarity: 24.12% Conservative: 2
Best Local Similarity: 23.81% Mismatches: 2
Query Match: 33.79% Indels: 732
DB: 21 Gaps: 4
US-10-075-846-4 (1-431) x AAC61681 (1-39796)

```

```

QY 21 ValLeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGln 40
DB 33015 CTCCTGTCAGAGGTGGCTTGGCAAAAGAGCAATCTGACACCAAGGGGTCCAG 33074
QY 41 PrometSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAla 60
DB 33075 CCCATGTCCTCCCTCGATTCCTAGACAAACTTATGGGGCAACATCTGATATGATGCC 33134
QY 61 ArgIleArgProAspPhe-Lys
DB 33135 AGGATTGCGCCCAATTTTAAGTAGAAGAAATCTCATCTATAAAACTCTCTCCCTCCC 33194
QY 67
DB 33195 ACCACCTTTGGCAGCACACTAGCCAGCCAGCCCTATTGCTTCCCTAAGGAAGAGGC 33254
QY 67
DB 33255 TEGAGGTCCAGAGGCTGGGGCTGGTTCTAACAGCTCAGAAGAGCCTCCTACCCCATTTGT 33314
QY 67
DB 33315 CCATGGCCTGTGTATCTGGACATGATTTCTCCACAGAGTGCCCTGACCTCAGCTCAGG 33374
QY 67
DB 33375 GAAGAGTTCTGGAGCTGCTCCAGCCCTCGAATGCGCTGCCAATGGCTGTGCTTT 33434
QY 68 -----GlyProProValAsnValThrcysAsnIlePheIleAsnSerPheSerSe 84
DB 33435 GCTACCTCAGAGCCCAACCGTGAAGTGCACGACATCTTCATCAACACAGTTTCAGCTC 33494
QY 84 rValThrLysThrThrMet
DB 33495 CATCACCAAGACCAACAAT-GGTAAGGATCTTCCTGCTCCCACTTCCAGCCTAGTGTA 33553
QY 90
DB 33554 GTGGAGAGAGCCCAACAGATTAACAGTGGCATCTTTGCCATATACAGCCCAAGAGAGT 33613
QY 90
DB 33614 TCCTTCCTCAGTGAATAATGCTTACTGCCCTGAGATGTGTCCCAACATTCCTCTCAT 33673
QY 90
DB 33674 GGCCCTCTGCCCTTACAGGCTTGTGGGCCCTTGGAATGGCAATGTTTCTGAAGGCC 33733
QY 91 -----AspTyrArgValAsnValPheLeuArgGlnGlnTyrPasnAs 104
DB 33734 CATATCTGACACCTCCAGAGACTACCGGGTGAATGCTCTTCGGCAACAGTGAATGA 33793
QY 104 pProArgLeuSerTyrArgGluTyrProAspAspSerIleAspLeuAspProSerMetLe 124
DB 33794 CCCAGCGCTGTCCTACGAGATATTCGATGACCTCTGGACCTCGATCCCTCATCTCT 33853
QY 124 uAspSerIleTyrPheProAspLeuPhePheIleAsnGluLysGlyValAsnPheHisG 144
DB 33854 GGACTCTATCTGGAGACCAAGACTTCTTTGCTATATAGAAAGGGCCCACTTCCATGA 33913
QY 144 uValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerI 164
DB 33914 GGTGACACGAGCAACAGTACTGCGCATCTTCAAGATGGGAATGTGCTACAGCAT 33973
QY 164 e-----
DB 33974 CAGGTGCACCGGTGATAGCCAGAGAGATTGCTTAAGGGAGAGAAATTTGGATGAAGA 34033
QY 164
DB 34034 CTGAGGGGTGGAGAGAGGTTCCCTTGACACTGTATGGGCCCAAGTAAGCCGATGTCACT 34093

```

OY	165	-----ArgLenthrLeuIIleuSerCysLeuMetLysA	178
Db	34094	TCCTTTCTACTGTCCTCCACTAGGCTGACCTTATTTTCTGCTCCCTGATGACCTCAAGA	34153
OY	178	snPhePromeLaspIleIntrhCysThrMetGlnLeuGluSer	192
Db	34154	ACITCCCATGGACATCCAGACCTGCAGATGCAGCTTGAGAGCTGAAGTGTCTATATA	34213
OY	192	-----	192
Db	34214	GAGTCACAGAGACCTGGAATAACAGCCAGTTAAACCCGACTCATGACATCAGACAGG	34273
OY	192	-----	192
Db	34274	AAAAGGCCACTCAGAGTTGGAGAGAACTTACTCAAGCACACCCAGCCCAATTTGGCAGC	34333
OY	192	-----	192
Db	34334	AGATGAGACTAGAAACCAGGGGCTCCTAACCCCTAGCCAGTGTTGTTTTCATCCCCACG	34393
OY	192	-----	192
Db	34394	TTATTGTGATCTGGCTGTTCATGAGGCTGGGAAAGGAAGCATCAGCCTTACGCA	34453
OY	192	-----	192
Db	34454	CAAAATGTTTGTGCTGTTTTGGGGGCGATTTGACGACATCTTCTGCCAAACGTA	34513
OY	192	-----	192
Db	34514	TGTCCCTCTCACAGTGCCTGGAAGTATGTTCCACCTTACCCTCTTGAAATTAAACATTTG	34573
OY	192	-----	192
Db	34574	AAGATGCTTATAAAATGCCAGAGTGGCAGGGTGGGGCAGCTGHTGGAGCCCTGATTCC	34633
OY	192	-----	192
Db	34634	CCACAAACATAGATTAACCTTTGTAGAGAGGTAGAGATGACAGTTCTAGGCAAGACAG	34693
OY	192	-----	192
Db	34694	TGCCTAGAAATGAAGAGGTGACCAAGGCTCTCTGTGGCAGGGTGCACAGCTGGAGGG	34753
OY	192	-----	192
Db	34754	GCTGCCAAAGTGCAGAGAAACCAAGATGACAAAGTTGGTTCATCATTTTCCACATC	34813
OY	192	-----	192
Db	34814	CCCTCCATCTCCTGGGTGAGAGAGACTGTTAGATCTGAGCCATCCACTAGATGTTCA	34873
OY	192	-----	192
Db	34874	CAGAGAAGCCTTATAAGCTGTTTCCAGCATATAAAATAACTGTGATGAATAGAAATGTCTC	34933
OY	192	-----	192
Db	34934	CTATGTGTCAGGCTTTAGAGTTCTTAAACCACTTTCACATCCAGAGCTCTCAATATCCC	34993
OY	192	-----	192
Db	34994	TTCATGGCAACGCTGAATAGTGGGAGAGACAGCTACTAATTTCCATTTTACAGAAAGG	35053
OY	192	-----	192
Db	35054	GAAGCGACTTACCCAAAGTCAAGAGCTACTTGGGTCAAGGCCCAACACTGGATGCTG	35113
OY	192	-----	192
Db	35114	TGAGCCTTGGTGTTCGTCCCATGCTACCTGACATGTGCAAAATATGGTCAATATAGCT	35173
OY	192	-----	192

Db	35174	GGCATGAGCCCTCATGAACTGCGATGATGATGTGGATTAACCTTGACATGTTTCAGACATTAG	35233
OY	192	-----	192
Db	35234	CTTTAAATTTGGCCCAACTCAAGGCCCTGGCTAGATCTATGGTATGACATTTGTCCAT	35293
OY	192	-----	192
Db	35294	GCATGACTGGCAGGATTAATATAGCTGTCTCTCCCTGAGGTGATGTGGCAGAGTCTTG	35353
OY	192	-----	192
Db	35354	GCTTGAAGGCTAGCTCAGCTTAAGACATGTGTTGGTACACATGCTTACTCCCTGGCT	35413
OY	192	-----	192
Db	35414	TGAGATATGGCAGGAGGCTCCAGCTGTGTTCCCTGACATCTTTCCACCCTTTGACCT	35473
OY	192	-----	192
Db	35474	GCTTTCTAGGTCGTGTAGAAATAATTCTCTTTGGGCTTTTCTCAGCTTGAGACAGGCT	35533
OY	192	-----	192
Db	35534	AGGAGAGAGCTGTATCTTCCGATTTACCTGCCACACAGAGTGGGCCATGAGACA	35593
OY	192	-----	192
Db	35594	ACTGACCAAGTCTCTCCAGAAAAGTTTCTTTCCATATAGACATGAACCTCTAA	35653
OY	192	-----	192
Db	35654	ACCCAGACCTTACTCCTCAGCTGTCTCTGCTGCTGTCTATGCTACCCCTACCCC	35713
OY	193	-----SerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrM	211
Db	35714	ACCCAGCATCATCTACTGTCAGGCCCTGTCATCTGTGTACATTGAGTGGCTACACCA	35773
OY	211	eLysAspLeuValPheGluTyrLeuLysAspLysProAlaValGlnValAlaGluGlyL	231
Db	35774	TGAAAGACCTCTGTTGGAGTGGCTGGACATGCTCTGCTGTCCAAGTGGCTGAGGGCC	35833
OY	231	eutThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyCysYThrLysHisT	251
Db	35834	TGACTCTGCCCATTTATCTTGGGGATGAGAAAGATCTAGGCTGTGTACCAACCACT	35893
OY	251	yrAsnThrGlyLys	255
Db	35894	ACAACACAGGTAAA	35907
RESULT 6			
AAD21397			
ID	AAD21397	standard; cDNA; 2400 BP.	
AC	AAD21397;		
XX			
DT	28-JAN-2002	(first entry)	
XX			
DE	R. sanguineus	glutamate-gated chloride channel 1 cDNA clone, T32.	
XX			
KW	Brown dog tick; glutamate-gated chloride channel; gluc11;		
KW	gluc12; crop protection; insecticide; nematocide; acaricide;		
KW	clone T32; ss.		
XX			
OS	Rhipicephalus sanguineus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	617..2170	
FT		/tag=a	
FT		/product="R. sanguineus gluc11 protein, T32"	
XX		/note="This region is specifically referred in claim 19"	

PN WO200174838-A1.  
 XX 11-OCT-2001..  
 XX 28-MAR-2001; 2001WO-US09905.  
 XX 31-MAR-2000; 2000US-193934P.  
 XX (MERI ) MERCK & CO INC.  
 PA Warmke JW, Yang Y, Cully DF, Hamelin MJ;  
 PI WPI: 2001-662963/76.  
 DR P-PSDB: MAE13039.  
 XX Novel L-glutamate-gated chloride channel proteins from Rhipicephalus  
 PT sanguineus for identifying compounds which modulate the channel  
 PT proteins, which are useful as insecticides, antelmintics and  
 PT acaricides .  
 PS Claim 18; Fig 5; 89pp; English.  
 XX The invention relates to Rhipicephalus sanguineus (brown dog tick)  
 CC L-glutamate-gated chloride channel proteins (Gluc11 and Gluc12)  
 CC and nucleic acid molecules encoding such proteins. Gluc1 channel  
 CC proteins are useful for identifying modulators. The compounds  
 CC identified as modulators are useful for insecticidal, mitacidal  
 CC and/or nematocidal treatment for use in animal and human health  
 CC and/or crop protection. The compounds are also useful in screening  
 CC for and selecting compounds active against parasitic invertebrate  
 CC species relevant to animal and human health, including worms,  
 CC fleas, ticks, mites and lice. Heterologous cell lines expressing  
 CC functional Gluc11 and Gluc12 channel functional forms are useful  
 CC for establishing functional or binding assays to identify novel  
 CC Gluc1 channel modulators. The present sequence is R. sanguineus  
 CC Gluc11 cDNA clone, T32.  
 XX  
 SO Sequence 2400 BP; 521 A; 769 C; 670 G; 440 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 7.47e-77 Length: 2400  
 Score: 754.50 Matches: 180  
 Percent Similarity: 55.21% Conservatave: 69  
 Best Local Similarity: 39.91% Mismatches: 133  
 Query Match: 33.50% Indels: 69  
 DB: Gaps: 22  
 US-10-075-846-4 (1-431) x AAD21397 (1-2400)  
 QY 5 ValPro--AlaThrLeuSerPheLeuLeuLeu-----Trp 15  
 Db 644 GTCCACGAGTGGCTGTAGCTTTTCTTGTGATTCCTGTCGTCCATCGGCAGCG 703  
 QY 16 -----ThrLeuPro-----GlyValValLeuLeuValAlaLeuAla 28  
 Db 704 GCCGAACGCTGCCTACGCCACCAACCCGTGGCCAGGGGGCGTCCGTCGCCGCCGCG 763  
 QY 29 -----LysGluGluValLysSerGlyThrLysGlySer 39  
 Db 764 ATGCTCCTGGGAAACAGCAAGTTCCTCCGCTACCAAGATTAAAGAGGC---AAGCAAAAT 820  
 QY 40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsp 59  
 Db 821 TTCCGCGCTATGAAGAACGCGATATTGGACAGCATCATGGCCAG---GTCGCTTATGAC 877  
 QY 60 AlaArgIleArgPro-----AsnPhelysGlyProProValAsnValThrcys 75  
 Db 878 TGCAGAGATCGGCCCATGGAATTAACACACAGACAGCGCGCTCTT---GTACGCGCTT 934  
 QY 76 AsnIlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsn 95  
 Db 935 AACATCTTTGTAGACAGTATGCGCAGATATGATGACGTCACCATGAGATACACAGTCA 994

QY 96 ValPheLeuArgGlnGlnIleTrpAsnAspProArgLeuSerTyrArgGluTyrProAspAsp 115  
 Db 995 ATGACGTTACAGAGAGCATGGCGGACGAGACTCCAGTACGACAGACTGGCGGCAG 1054  
 QY 116 SerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAla 135  
 Db 1055 GTTCGCTACCTCAGACGCTCAGCAACCGGACAGGCTTTGGAGCGCGACTGTTTCTCC 1114  
 QY 136 AsnGluLysGlyAlaAsnPheHisGluValThrThrAspAsnLysLeuArgIlePhe 155  
 Db 1115 AACGGAAGAGGAGGACACTTCCACACATCATTCGCCCAACGTCCTTCTACGATCAT 1174  
 QY 156 LysAsnGluAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAsp 175  
 Db 1175 CCCAACGGCGAGCTTCCTTCATGATCAGATATTCCTTGGCTTCATGTCGATGAC 1234  
 QY 176 LeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIle 195  
 Db 1235 CTGAATTTTATCCTTGGATTAACAAATCTGCTCTATC----- 1273  
 QY 196 LeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuVal 215  
 Db 1274 -----GTCATGGTGAAGCTATGGGTATACACAGAGACTGGTG 1312  
 QY 216 PheGluTrpLeuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGln 235  
 Db 1313 TTTCTATGTAAGAAAGGGGAGATCT---GTACAGGTCAACAAATTCCTCACTGCCAGT 1369  
 QY 236 PheIleLeu---ArgAspGluLysAspLeuGlyCysThrLysHisTyrAsnThrGly 254  
 Db 1370 TTCACGCTGGAAGAGTTTCAACCGACTAC-----TGCACACCTCGGACCAACACGCGC 1423  
 QY 255 LysPheThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIle 274  
 Db 1424 GAGTACACTGCTCTGCGGCTGAGCTGCTTCAAGCCGAGTTCAGCTACTACTGATC 1483  
 QY 275 GlnMetYrIleProSerLeuLeuValIleLeuSerTrpValSerPheTrpIleAsn 294  
 Db 1484 CAGATCTACATCCGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1543  
 QY 295 MetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrThr 314  
 Db 1544 CCCACCTCGATCCGCGCGGAGTGTGCTGGGCTGCACACCCGCTCACCATTGCGCAGC 1603  
 QY 315 GlnSerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIle 334  
 Db 1604 CAGATATCGGCGCATCAACGCTCGCTCCCTTCTTCTACACCAAGCCATTGAGCTG 1663  
 QY 335 TrpMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsn 354  
 Db 1664 TGGACCGCGCTGCTGCTGATCTGATTTGCGCGCTCCTCGAGTTGCGCCCTGCTCAAC 1723  
 QY 355 PheValSerArgGlnHisLysGluPheIleArgLeuAlaGlyArgGlnArgGln--- 373  
 Db 1724 TACGCTCGCGGTAGATTCACGCCGCGCAACATGCAAGAGCAGAAAGCAAGAAATGG 1783  
 QY 374 -----ArgLeuGluGluAspIleIleGlnGlnSerArgPheTyrPheArgGly 389  
 Db 1784 GAGCTCGAGCGCCCTTGATCGGACCATCGGACAGAGCGCGCACCATTTGCCATG 1843  
 QY 390 TyrGlyLeuGlyHisCysLeuGlnAlaArgAspGlyProMetGluGlySerGlyIle 409  
 Db 1844 GTGAGCTCCGCGC-----ProGlnProProAlaPro 417  
 QY 410 TyrSer-----ProGlnProProAlaPro 417  
 Db 1871 ATGGCGGCAACCTGCGCACACCGCGCTGCGC 1903  
 RESULT 7  
 ID AAD26939 standard; cdna: 4621 BP.  
 XX AAD26939;  
 AC AAD26939;



```

Db      1292 CAAGATCGCGCAGATCGAGATCATCATCCCGCCG 1327
RESULT 8
ID      AAD26937
XX      AAD26937 standard; DNA; 1609 BP.
XX      AAD26937;
XX      09-APR-2002 (first entry)
XX      Heliothis sp. HEG3E(4)-2 plasmid DNA.
XX      Lepidopteran glutamate-gated chloride channel; insecticide;
XX      HEG3E(4)-2 plasmid; cyclic; circular; ds.
XX      Heliothis sp.
OS      Heliothis sp.
Key      Location/Qualifiers
FT      misc-feature 1 /tag= a
FT      /note= "Base C is present in the sequence shown in
FT      the sequence listing"
PN      US6329174-B1.
PD      11-DEC-2001.
XX      13-JUN-2000; 2000US-0592891.
XX      13-JUN-2000; 2000US-0592891.
XX      13-JUN-2000; 2000US-0592891.
XX      (AVET ) AVENTIS CROPS SCIENCE SA.
XX      Wang XM, Sarda XG, Tomalski MD, Wingate VPM;
XX      WPI: 2002-121133/16.
XX      New nucleic acid encoding lepidopteran chloride channel, useful for
XX      screening agents for insecticidal activity -
XX      Example 1; Column 9-10; 18pp; English.
XX      The invention relates to nucleic acid encoding lepidopteran
XX      glutamate-gated chloride channel. Glutamate-gated chloride channels
XX      are a family of ligand-gated chloride channels unique to invertebrates.
XX      The DNA of the invention is used for recombinant production of
XX      lepidopteran glutamate-gated chloride channel and this is used,
XX      optionally in the form of membrane preparations or recombinant cells,
XX      in specific-binding or functional assays for identifying potential
XX      insecticides. The present sequence is Heliothis sp. HEG3E(4)-2 plasmid
XX      DNA used in the exemplification of the invention.
XX      Note: This sequence SEQ.ID.NO.11 is stated to be similar to the
XX      sequence shown in sequence listing of the specification. However
XX      these sequences differ.
XX      Sequence 1609 BP; 411 A; 410 C; 393 G; 395 T; 0 other;
XX
Alignment Scores:
Pred. No.: 1.75e-76 Length: 1609
Score: 749.00 Matches: 158
Percent Similarity: 61.568 Conservative: 55
Best Local Similarity: 45.668 Mismatches: 109
Query Match: 33.268 Indels: 24
DB: 24 Gaps: 7
US-10-075-846-4 (1-431) x AAD26937 (1-1609)
XX
31 GUVALYLSERGXTHTLYSGLYSERGINPROMETSERPROSERASPHENEAUSPLY 50
DB      211 GAATGCAAGAACGGGCGGAGAAATCACTTCCGAGAGAAAGAGAGATCCCTGGATCAG 270
XX      51 LEUWETGTYATGTTHTSERGLTYTTRASPAALATGIIATGPR-----ASNPHELYS 67

```

Db	271	ATCCTGGGC---CCCGGAGGTACGACCCAGATCTACAGACCTCGGGGATATCAAGCACT	327	111
Qy	68	GLYProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLys	87	111
Db	328	GATGGCCACGCGGAGTAGAGCGCATATATTTTCCGAAGTATATCAAGATCGATGAC	387	111
Qy	88	ThrThrMetAspTyrArgValAsnValPheLeuArgIleGlnIntrpAsnAspProArgLeu	107	111
Db	388	GTCACAAATGAAATCTCCGTAACATTTAAACGTTTGGGAAACAAATGGTTATGAACGGCTC	447	111
Qy	108	SerTyrArgIleuTyrProAspAspSerLeuAsnAspLeuAspSerMetLeuAspSerIle	127	111
Db	448	AAATTCAATATCTTGGAGCTCGCCTAAATACTCTGACACTGACTGAACCCACAGAGTC	507	111
Qy	128	TrpLysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheIleGlnValThrThr	147	111
Db	508	TGGATGCCGTGATCTATTTCTTCCACAGAGAAGAGGTGATTTCACAAATCATCATG	567	111
Qy	148	AspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThr	167	111
Db	568	CCGAAAGTGATCAATCCGAAATCTTCCCAACGCGCAACGTGGTACACATCCGAATCTCC	627	111
Qy	168	LeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnIntrpCysThr	187	111
Db	628	CTGACGCTCTGTGCCCCCATTAACCTCAAGTTGTACCCCCGTGGATTAAGCAAGCTGCTG	687	111
Qy	188	MetGlnLeuGlnSerSerSerIleLeuLysSerProLeuProSerLeuSerLeuVal	207	111
Db	688	CTCAGAGAG-----GCTAGTTAT	705	111
Qy	208	GlyTyrThrMetLysAspLeuValPheGluTrpLeuGlnAspAlaProAlaValGlnVal	227	111
Db	706	GTTGGAGCACAGACGACTTGTGTCTCTTAAGAAAGGCGACCGCG---GTGCAGAGTG	762	111
Qy	228	AlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeu---GlyCys	246	111
Db	763	GTGAAAACCTTACACCTGCTCGGTCTACCGCTG----GAGAAGTTCTCTACTGACTAC	816	111
Qy	247	CysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGlu	266	111
Db	817	TGCAACAGTAGACTAATAACCGGTGAATACAGTTGCCGTAAGGTAGACCTGCTTCAMA	876	111
Qy	267	ArgIleMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeu	286	111
Db	877	CGCAGATTCAGTTACTACTGATCCAGATCTCAATTCCTCGTCTGATGCTGTCATCGTG	936	111
Qy	287	SerTrpValSerPheTrpIleAsnMetAspAlaProAlaArgValGlyLeuGlyIle	306	111
Db	937	TCTGGGGTCTCTGTGGCTGGACAGGAGCTGCTCGTGAGAGGTCTCACTAGAGGTG	996	111
Qy	307	ThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVal	326	111
Db	997	ACGACTTACTTACATGCGCAGCCAGTCGTCAGAGCATCAACGGGTCCCTACACCGGTG	1056	111
Qy	327	SerTyrValLysAlaIleAspIleTrpMetAlaValLysLeuLeuPheValPheAlaAla	346	111
Db	1057	TCTCACACGAAGCATGATGATGTGGACTGGATTATCTCAATTCGTAATTCGGAGCG	1116	111
Qy	347	LeuLeuGluTyrTrpAlaAlaIleAsnPheValSerArg-----GlnHisLysGluPheIle	364	111
Db	1117	CTACTAGAGTTGGCTGCTGCAACTATGCGTCTGCTGACATGACACCGAGACATCG	1176	111
Qy	365	ArgLeuArgArgArgGln	370	111
Db	1177	AAGAAAGCGAAGACGGGAG	1194	111
RESULT 9				
AAD21378				
XX	AAD21378	standard; DNA: 1197 BP.		
XX	AAD21378;			



XX R. sanguineus GluCl DNA probe to clone GABA-gated chloride channel DNA.  
 DE Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;  
 KW GABA-gated chloride channel; recombinant expression; domestic animal;  
 KW GluCl DNA; probe; ss.  
 OS Rhipicephalus sanguineus.  
 XX MO200174884-A1.  
 XX 11-OCT-2001.  
 XX 28-MAR-2001; 2001WO-US09955.  
 XX 31-MAR-2000; 2000US-193791P.  
 XX (MERI ) MERCK & CO INC.  
 XX Zheng Y, Cully D, Ludmerer S;  
 DR WPI: 2002-010778/01.  
 XX New polypeptide useful for preventing or treating tick infestation, in  
 PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
 PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid  
 PT (GABA)-gated chloride channel -  
 XX  
 XX Example 1; Page 23-24; 59pp: English.  
 CC The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
 CC channels and their corresponding nucleic acid molecules. GABA-gated  
 CC chloride channel proteins and DNA's are useful for preventing and  
 CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
 CC deer, or other wild or domesticated animals. The nucleic acids are useful  
 CC as hybridisation probes or Polymerase Chain Reaction primers for  
 CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
 CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
 CC acids are also useful for the recombinant expression of D. variabilis  
 CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
 CC proteins exert toxic effects on other ticks or related parasites such as  
 CC mites. The present sequence is Rhipicephalus sanguineus GluCl DNA probe  
 CC which is used for cloning Dermacentor variabilis GABA-gated chloride  
 CC channel DNA.  
 XX  
 XX Sequence 1197 BP; 268 A; 358 C; 311 G; 260 T; 0 other:  
 SO  
 Alignment Scores:  
 Score: 6.29e-76 Length: 1197  
 Percent Similarity: 742.50 Matches: 157  
 Best Local Similarity: 61.06% Conservative: 61  
 Query Match: 43.98% Mismatches: 110  
 DB: 24 Gaps: 8  
 US-10-075-846-4 (1-431) x AAD22082 (1-1197)  
 OY 48 LeuaspSerIleTrrPlysProAspLeuPheAlaAsnGluYsGlyAlaAsnPhetHis 64  
 DB 7 TTGGACAGCATCATGTGGCCAG--GTCGTATATGACAGATCCGCGCATGGGAATT 63  
 OY 65 ---AsnPhelysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83  
 DB 64 AACAAACACAGAGCGCGGCTCTT--GTACCGCTTAACATCTTGTAAAGAATATCGGC 120  
 OY 84 SerValThrLysThrThMetAspTyrArgValAsnValPheLeuArgGlnGlnTrrPasn 103  
 DB 121 AGAFTTGTATGACGTCACCATGAGTACACAGTGCATAAGACAGCTTCACAGACAGCGCG 180  
 OY 104 AspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMet 123  
 DB 181 GACGAGACATCTCCAGTACGACGACTTGGCGCGCAGTTCCTACCTGACGCTACCGAA 240

OY 124 LeuaspSerIleTrrPlysProAspLeuPheAlaAsnGluYsGlyAlaAsnPhetHis 143  
 DB 241 CCGGACAGCATTTTGAACCCGAGCTGTTTCTTCCAAAGAAAGAGGACACTTCAC 300  
 OY 144 GluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSer 163  
 DB 301 AACATCATCATGCCAACGCTTCTACGCAATACATCCAGCGAGCGGCTTCTTACGC 360  
 OY 164 IleArgLeuThrLysIleLeuSerCysLeuMetAspLeuLysAsnPhetAspIle 183  
 DB 361 ATCAGAAATATCTTGTCTGCTTCTTCAATGCCATGACCGAATATTTATCTTGGATAA 420  
 OY 184 GlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeuProSerLeu 203  
 DB 421 CAAATCTGCTTATC-----GTC 438  
 OY 204 SerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrrPleuGluAspAlaPro 223  
 DB 439 ATGTGAGCTATGGGTATACACAGAGGACCTGGTGTCTTATGAAAGAGGCGGATCT 498  
 OY 224 AlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeu---ArgAspGluYs 242  
 DB 499 ---GTACAGTCACAAATAATCTCCACTTGGCCAGCTTTCACGCTGGAAGGCTTTCANAC 555  
 OY 243 AspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLys 262  
 DB 556 GACTAC-----TGCACAGTCGAGCACACCTGGCGAGTACAGCTGTCGCGGTGAC 609  
 OY 263 PheHisLeuGluArgGlnMetGlyTyrTyrThrLeuIleGlnMetTyrIleProSerLeuLeu 282  
 DB 610 CTGCTGTTCAAGCGGAGATTGACTACTGATCTGATCCAGATCTTCAATCCCTGCTGCATG 669  
 OY 283 IleValIleLeuSerTrrPvalSerPheThrPheLeuMetAspAlaAlaProAlaArgVal 302  
 DB 670 CTGTCATCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729  
 OY 303 GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSer 322  
 DB 730 TCGCTGGCGGTCACACCTCTCCTACATGGCCAGATATGCGGATCATCAACGCTCG 789  
 OY 323 LeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCysLeuPhe 342  
 DB 790 CTGCTCCCGCTTCTTCTACACCAAGGCCATTACGTGTGAGACGGCGTCTGTACCTTC 849  
 OY 343 ValPheAlaLeuLeuGluTyrAlaIleAlaIleAsnPheValSerArgGlnHisLysGlu 362  
 DB 850 GTATTGCGGCGGCTCTCGATGCTGCCCTGCTCAATGACGCTCGGCTGATTCACGC 909  
 OY 363 PheIleArgLeuArgArgArgGln-----ArgLeuGlnGlu 377  
 DB 910 CGGACAGACATGCAAGAACAGAACAGAAATGGAGAGCTCGAGCCGCCCTCTGACTCG 969  
 OY 378 AspIleIleGlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlnHis 394  
 DB 970 GACCACCTGGAAGAGCGGCGCCACCAAGCTTCCATGAGGCGCGTGTGTCAC 1020  
 RESULT 11  
 AAD21395 standard; CDNA: 2138 BP.  
 ID AAD21395;  
 AC AAD21395;  
 XX 28-JAN-2002 (first entry)  
 DE R. sanguineus glutamate-gated chloride channel 1 cDNA clone, T12.  
 XX Brown dog tick; glutamate-gated chloride channel; GluCl1;  
 KW GluCl2, crop protection; insecticide; nematocide; acaricide;  
 KW clone T12; ss.  
 OS Rhipicephalus sanguineus.  
 XX  
 FH Key Location/Qualifiers



FT	CDS		331..1683	
FT			/+tag= a	
FT			/product= "R. sanguineus Gluc11 protein, T12"	
XX			/note= "this region is specifically referred in claim 7"	
XX				
PN			MO200174838-A1.	
XX				
PD			11-OCT-2001.	
XX				
XX			28-MAR-2001; 2001WO-US09905.	
PR				
PP			31-MAR-2000; 2000US-193934P.	
XX				
PA			(MERI ) MERCK & CO INC.	
XX				
PI			Warmke JW, Yang Y, Cully DF, Hameljin MJ;	
XX				
DR			WPI: 2001-662963/76.	
XX			P-PSDB: AAE13037.	
XX				
PT			Novel L-glutamate-gated chloride channel proteins from Rhipicephalus	
PT			sanguineus for identifying compounds which modulate the channel	
PT			proteins, which are useful as insecticides, anthelmintics and	
PT			acaricides	
XX				
PS			Claim 6; Fig 1; 89pp; English.	
XX				
CC			The invention relates to Rhipicephalus sanguineus (brown dog tick)	
CC			L-glutamate-gated chloride channel proteins (Gluc11 and Gluc12)	
CC			and nucleic acid molecules encoding such proteins. Gluc1 channel	
CC			proteins are useful for identifying modulators. The compounds	
CC			identified as modulators are useful for insecticidal, mitacidal	
CC			and/or nematocidal treatment for use in animal and human health	
CC			and/or crop protection. The compounds are also useful in screening	
CC			for and selecting compounds active against parasitic invertebrate	
CC			species relevant to animal and human health, including worms,	
CC			fleas, ticks, mites and lice. Heterologous cell lines expressing	
CC			functional Gluc11 and Gluc12 channel functional forms are useful	
CC			for establishing functional or binding assays to identify novel	
CC			Gluc1 channel modulators. The present sequence is R. sanguineus	
CC			Gluc11 cDNA clone, T12.	
XX				
XX				
S0			Sequence 2138 BP; 502 A; 661 C; 561 G; 414 T; 0 other:	
Alignment Scores:				
Pred. NO.:	1,55e-75	Length:	2138	
Score:	742.50	Matches:	157	
Percent Similarity:	61.06%	Conservative:	61	
Best Local Similarity:	43.98%	Mismatches:	110	
Query Match:	32.97%	Indels:	29	
DB:	22	Gaps:	8	
US-10-075-846-4 (1-431) x AAD21395 (1-2138)				
OY	48	LeuasplylsUwNeIgVArGThrSerGlyTrAsPaLaArgLIeaRGPro-----	64	
Dd	454	TTCGACAGCATCATTATGGCCAG---GOTCGTATGACTGCAGAGATCCGCCCATGGAAAT	510	
OY	65	---AsmPhelysGIyPrOProVaLAsnValImrCYcSAnIIIEHeIIIEAsnSerPheser	83	
Dd	511	AACAACACAGACGGCGCGCTCT--GTACCCTTAACATCTTTGAAGAAGATATCGGC	567	
OY	84	SeGVAlnHrlsrThrMeIAspTYrAryVAIsnValPheLeuArgGLInTPAsn	103	
Dd	568	AGAAITGATCACTCACCATCGAGTACACAGTGCAAATGACGTTCCAGAGACATGGCGG	622	
OY	104	ASpPrOArgLeuSerTYrArGluTyrrProAsPasPsSerLeuAspLeuAspProSerMet	123	
Dd	628	GACAGACAGACTCCAGTACAGACACACTGGCGGCCAGTTCCGTACTCATGCCATCACC	687	
OY	124	LeuAspSerIIErrIPryrProAsPLeuPhePheAlaAsnGluVsgLYAlAsnPheHis	143	
Dd	688	CCGGACAAAGCTTTGGAGCGGACCGCTTTTCTCCAACAGAAAGAGGAGCACTTCAC	747	

Oy		144	GluValThrThrAspAsnLysIleLeuArgTllePheLysasnGlyAsnValIleuTySer	163
Dd		748	AACATCATCATGCCCAAGCTGCTTTCTACGCATTACATCCCAAGCGCAGCTTCTTGAC	807
Oy		164	IleArgLeuThrIleuIleLeuSerCysLeuMetaspLeuLysAsnProMetAspIle	183
Dd		808	ATCAGAATAATTCCTGGTGTCTTCTATGTCCATGAACCTGAAAATTTATCTTGATAA	867
Oy		184	GlnThrcysThrMetGlnLeuGlnSerSerIleLeucysSerProLeuProSerLeu	203
Dd		868	CAAAATCTGCTTATC-----GTC	885
Oy		204	SerLeuSerValIGlYTrThrMetLysAspLeuAlPhegluTrPleuGluAspAlaPro	223
Dd		886	ATGGTGACTATGGGTATACACAAGAGACTGGTGTGTTCTAATGAAAAGGGGATCT	945
Oy		224	AlaValGlnAlaIaGlugLYLeuThrLeuProGlnPheIleLeu---ArgAspGluLys	242
Dd		946	--GTACAGGTGCACAAAAATCTCCACTGTCGCAGTTTCAACGCTGGAAAGTTTCAAAC	1002
Oy		243	AspLeuGIcysCysThrLysHisTYrAsnThrGLylusPheThrCysIleGluValLys	262
Dd		1003	GACATAC-----TTCACCAGTCGGAGCACAACTGGCGATACACTCTCTGGCGGTGAC	1056
Oy		263	PheHisLeuGluArgrGlnMetGLYTyrTYrLeuIleGlnMetTYrIleProSerLeuLeu	282
Dd		1057	CTGGTGTTCAAAGCCGAGTTACGTAAGTACTACCTGATCCAGATCTACATCCCTGCTCATG	1116
Oy		283	IleValIleLeuSerTYrValISerPheTrPllaasmetaspAlaAlaProAlaIargVal	302
Dd		1117	CTGGTCAATCGTGTCTGGGTGCTGCTGCTGACCCCACTCGATCCGCGCGAGATG	1176
Oy		303	GlyLeuGIcylIeThrThValLeuThrMetThrThrGlnInserserylSerArgAlaSer	322
Dd		1177	TTCGTGGCGGTACACCCTGCTCACTACATGGCCACGAGATATGGGCGATCAACGCTCG	1236
Oy		323	LeuProLysValISerTYrValLysAlaIleAspIleTrpmelaIalALysLeuPhe	342
Dd		1237	CTGGCTCCCGTTCTCTACACCAAGCCATTACGTTGGACCGGCGTGTGTCACCTTC	1296
Oy		343	ValPheAlaIaleuLeuGluTYrAlaAlaIleasnPheValISerArgGlnHisLysGlu	362
Dd		1297	GTATTGCGGCGGCTCTGAGTAGTTCGGCTGTGTAATAAGCGCTCGGCTCGATTCACGC	1356
Oy		363	PheIleArgLeuArGatArgGlnArgGln-----ArgLeuGluGlu	377
Dd		1357	CGGCAAGAACATGACAGAACACAGACAGAAATGGAGACTCGAGCGCCCTGTGACTCG	1416
Oy		378	AspIleIleGlnIuSerArgRphetyrPheArgGlyTYrGlyLeuGlnHis	394
Dd		1417	GACCACCTGGAGAGAGCGGCCACCACTGTCGCATGAGCGCGCTGTGTGCAC	1467
RESULT 12				
ID	AAD21396		standard; cDNA: 2289 BP.	
XX	AC	AAD21396;		
XX	XX	28-JAN-2002	(first entry)	
DE	XX	R. sanguineus	glutamate-gated chloride channel 1 cDNA clone, T82.	
KW	XX	Brown dog tick:	glutamate-gated chloride channel; GluCl1.	
KW	XX	GluCl2:	crop protection; insecticide; nematocide; acaricide;	
KW	XX	clone T82;	ss.	
OS	XX	Rhipicephalus	sanguineus.	
FH	Key	Location/Qualifiers		
FT	CDS	502..1854	/tag= a	
FT	FT	/product= "R. sanguineus	GluCl1 protein, T82"	

FT /note= "This region is specifically referred in claim 13"

XX WO200174838-A1.

XX 11-OCT-2001.

XX 28-MAR-2001; 2001WO-US09905.

XX 31-MAR-2000; 2000US-193934P.

XX (MERI ) MERCK & CO INC.

XX Warmke JW, Yang Y, Cully DF, Hamelin MJ;

XX WPI; 2001-662963/76.

XX P-PSDB; AAE13038.

XX Novel L-glutamate-gated chloride channel proteins from Rhipicephalus

PT sanguineus for identifying compounds which modulate the channel

PT proteins, which are useful as insecticides, antheimintics and

PT acaricides

XX

XX Claim 12; Fig 3; 89pp; English.

XX

XX The invention relates to Rhipicephalus sanguineus (brown dog tick)

CC L-glutamate-gated chloride channel proteins (Gluc11 and Gluc12)

CC and nucleic acid molecules encoding such proteins. Gluc1 channel

CC proteins are useful for identifying modulators. The compounds

CC identified as modulators are useful for insecticidal, mitacidal

CC and/or nematocidal treatment for use in animal and human health

CC and/or crop protection. The compounds are also useful in screening

CC for and selecting compounds active against parasitic invertebrate

CC species relevant to animal and human health, including worms,

CC fleas, ticks, mites and lice. Heterologous cell lines expressing

CC functional Gluc11 and Gluc12 channel functional forms are useful

CC for establishing functional or binding assays to identify novel

CC Gluc1 channel modulators. The present sequence is R. sanguineus

CC Gluc11 cDNA clone, 782.

XX

SQ Sequence 2289 BP; 509 A; 727 C; 609 G; 444 T; 0 other;

Alignment Scores:

Pred. No.: 1 728-75 Length: 2289

Score: 742.50 Matches: 157

Percent Similarity: 61.06% Conservative: 61

Best Local Similarity: 43.98% Mismatches: 110

Query Match: 32.97% Indels: 29

DB: 22 Gaps: 8

US-10-075-846-4 (1-431) x AAD21396 (1-2289)

QY 48 LeuAspLysLeuMetGlyArgThrSerGlyTrpAspAlaArgLeuArgPro----- 64

DB 625 TTGGACACATCATGTGGCCAG---GGTGGTATATGCTCGAGATCCGGCCCATGGGAAAT 681

QY 65 ---AsnPhelysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83

DB 682 AACACACAGACGAGCGCCGCTCTT---GTACGCGTAACTTGTAAAGATATCGGC 738

QY 84 SerValThrLysThrThMetAspPyrArgValAsnValPheLeuArgGlnGlnTrpAsn 103

DB 739 AGAATTTGATGACGTCCACATGAGTACAGAGTCAAAAGAGCTTCAGAGAGAGCGCG 798

QY 104 AspProArgLeuSerTyrArgGlyLysTrpAspAspSerLeuAspProSerMet 123

DB 799 GACGAGACACTTCATGACGACACTTGGCGCCAGGTTCCGTACCTACGCTCACCA 858

QY 124 LeuAspSerIleTyrPheProAspLeuPhePheAlaAsnGlyLysGlyAlaAsnPhenHis 143

DB 859 CCGGACAGACTTTTGAACCGGACCTGTTTTCACAGAAAGAGGACACTTCAC 918

QY 144 GluValThrThrAspAsnLysLeuArgIlePheLysAsnGlyAsnValLeuTyrSer 163

DB 919 AACATCATCATGCCAACGCTCTTCTACGATACATCCCAAGCGGAGCTTCTTACGC 978

QY 164 IleArgLeuThrLeuIleLeuSerCysLeuMetAspLysAsnPhenPheMetAspIle 183

DB 979 ATCAGAAATATCTTGGTCTTTCATGTCGATGACCACTGAAATTTTATCTTGGATAA 1038

QY 184 GlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeuProSerLeu 203

DB 1039 CAAATCTCTCTATC-----GTC 1056

QY 204 SerLeuSerValGlyTyrThrMetLysAspLeuValPheGlyLysLeuAspAlaPro 223

DB 1057 ATGGAGCTATGGGTATACACAGAGGACCTGGTGTCTTATGAAAGAGGCGATCT 1116

QY 224 AlaValGlnValAlaGlyLeuThrLeuProGlnPheIleLeu---ArgAspGlyLys 242

DB 1117 ---GTACAGGTACAAAATCTCCACTTGGCAGCTTTCAGCGTGGAAAGTTTCAAAAC 1173

QY 243 AspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlyValLys 262

DB 1174 GACTAC-----TGACACAGTCGAGCAACAGCTGGAGTACAGCTGCTTCCGCGTGAC 1227

QY 263 PheHisLeuGluAlaArgLysMetGlyTyrTyrLeuLeuIleGlnMetTyrIleProSerLeu 282

DB 1228 CTGGTGTTCACAGCGGAGTTCAGCTACTACCTGATCCAGATCTCATCCCTGCTGATG 1287

QY 283 IleValIleLeuSerTyrPheValSerPheThrIleAsnMetAspAlaAlaProAlaArgVal 302

DB 1288 CTGGCATCTGTCTCTGGGTCTCTGTGCTGCATCCACCCTGATGCCGCGGAGATG 1347

QY 303 GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSer 322

DB 1348 TCGCGGCGCGACACACCTGCTCAGATGCGACAGATATGCGGATGACAGCGCTCG 1407

QY 323 LeuProLysValSerTyrValLysAlaIleAspIleThrPheMetAlaValCysLeuPhe 342

DB 1408 CTGCTCCGCTTCTTACACCAAGGCAATGACGTTGAGCGGCGCTGTCTGATCCCTC 1467

QY 343 ValPheAlaAlaLeuLeuGlyTyrAlaIleAsnPheValSerArgGlnHisLysGlu 362

DB 1468 GTATTCGCGCGCTCTCTGATGCTGCGCTGCAACTGAGCTGCGGCTGAGATTCACGC 1527

QY 363 PheIleArgLeuAlaArgArgGlnArgGln-----ArgLeuGlu 377

DB 1528 CGCGAAGACATGACAAACAGACAGAGGAAATGGAGCTCGAGCCGCTGAGCTCG 1587

QY 378 AspIleIleGlnGluSerArgPheThrPheArgGlyTyrGlyLeuGlyHis 394

DB 1588 GACCACTGAGAGGAGCGCGCCACCACTTCCCATGAGGCGCGTGTGTGCAC 1638

RESULT 13

AAD22072

ID AAD22072 standard; DNA; 1614 BP.

XX

AC AAD22072;

XX

DE 12-FEB-2002 (first entry)

XX

DE Dermacentor variabilis clone 5 GABA-gated chloride channel DNA.

KW Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;

KW GABA-gated chloride channel; recombinant expression; domestic animal; ds.

OS Dermacentor variabilis.

XX

XX

XX key Location/Qualifiers

FT 1..1614

FT /tag= a

FT /product= "GABA-gated chloride channel protein"

PN WO200174884-A1.

PD 11-OCT-2001.

PF	28-MAR-2001: 2001MO-US09955.	
XX		
XX	31-MAR-2000: 2000US-193791P.	
XX		
XX	(MERI ) MERCK & CO INC.	
XX		
PI	Zheng Y, Cully D, Luderer S;	
XX		
DR	WPI: 2002-010778/01.	
XX	P-PSDB: AAE13314.	
PT	New polypeptide useful for preventing or treating tick infestation, in	
PT	animals, dogs, cattle, horses, deer, or other wild or domesticated	
PT	animals, comprises the Dermacentor variabilis gamma-aminobutyric acid	
PT	(GABA)-gated chloride channel -	
XX		
PS	Claim 8: Fig 3: 59pp: English.	
XX		
CC	The invention relates to gamma-aminobutyric acid (GABA)-gated chloride	
CC	channels and their corresponding nucleic acid molecules. GABA-gated	
CC	chloride channel proteins and DNA's are useful for preventing and	
CC	treating tick infestation, particularly in humans, dogs, cattle, horses,	
CC	deer, or other wild or domesticated animals. The nucleic acids are useful	
CC	as hybridisation probes or Polymerase Chain Reaction primers for	
CC	identifying the presence of Dermacentor variabilis GABA-gated chloride	
CC	channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic	
CC	acids are also useful for the recombinant expression of D. variabilis	
CC	GABA-gated chloride channel proteins. GABA-gated chloride channel	
CC	proteins exert toxic effects on other ticks or related parasites such as	
CC	mites. The present sequence is Dermacentor variabilis clone 5 GABA-gated	
CC	chloride channel DNA.	
XX		
SO	Sequence 1614 BP; 362 A; 475 C; 429 G; 348 T; 0 other:	
	Alignment Scores:	
	Pred. NO: 3,81e-75 Length: 1614	
	Score: 737.50 Matches: 164	
	Percent Similarity: 58.94% Conservative: 70	
	Best local Similarity: 41.31% Mismatches: 107	
	Query Match: 32.75% Indels: 57	
	Db: 24 Gaps: 13	
	US-10-075-846-4 (1-431) x AAD22072 (1-1614)	
OY	55 TSSerGLTYTyrAspLarArgLleArgProAsnPhelySGlyProProValAsnValThr 74	
Db	154 ACACGCGTGGAGCACGAGAGGCTGAGGCCAAATTATGGCGCGTCCAGTGGAGGTTGGC 213	
OY	75 CysAsnIlePheIleAsnSerPheSerSerValThrIlyStrThrMetLasPTyrArgVal 94	
Db	214 GTCACATAGCAGATTTATTCAGCAATAAGTACAGTCTCGAAGTAAACAAATGAGACTTACTTC 273	
OY	95 AsnValPheLeuArgGlnGlnIrrpAsnSpproArgLeuSerTYrArgGluTYrProAsp 114	
Db	274 GACCTTATTATTCGCGCAATCGTGGCGGCGAGCAGCACTCTCGTTCCAGAAAAAGCCAGAC 333	
OY	115 ---AspSerIleuAspIleuAspProserMetIleuAspSerIleTPlySProAspLeuPhe 133	
Db	334 CTGCAAGACATGACTGTGGCCGCTGAAGTGGCCGAGAGGATCTGGTAACTCCGACACCTTC 393	
OY	134 PheAlaAsnGluLysGluValaAsnPhenIleGluValaThrThrAspAsnLysLeuLeuArg 153	
Db	394 TTGCGCCAGCAGAAAGAGCCCTTACTTTCATGCGGCCACCAACAGCCCAACTTCTCCGCG 453	
OY	154 IlePheLysAsnGluValLeuTYrSerIleArgLeuThrIleuLleLeuSerCysLeu 173	
Db	454 ATGGCGCTCCGAGAGAGAGGTTTTCGCAATATTCACACGAGGAGACTCCGCGCTGCCCA 513	
OY	174 MetAspIleuLysAsnPheProMetLasPliGlnThrCysThrMetGlnLeuGluSerSer 193	
Db	514 ATGATATCTCGATATCTTCCCGATGACACACACAGCGTCTCACTATATAGATAGAA----- 567	

OY	194	SerIleLeuCysSerProIeuProSerIeuSerLeuSerValGlyThrMetLysASP	213
Db	568	-----AGCTTTGGTATACCATGAAGAC	591
OY	214	LeuValPheGluTrpLeuGluAspAlaProAlaValGlnValAlaGluIleuThrLeu	233
Db	592	ATCCGCTACCGGTGGTGGAGGATGACACCTCCGTCGGATCGCCAAAGAGGTAAGATTG	651
OY	234	ProGlnPhe-----IleLeuArgAspGluLysAspLeuIleCysCysThrLys	249
Db	652	CCGCGAGTTCAAGTCCCTCGCTCGACGCCAAAAGCCAAAAGAGATTGCCCTAACGACAGGA	711
OY	250	HisTrpAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGlnMet	269
Db	712	AACTACTCC-----CGCTGGTAGT---GAATAACGGGTTCC-----GCCCGCTCCATG	756
OY	270	GlyTrpTrpLeuIleGlnMetTrpIleProSerLeuLeuIleValIleLeuSerTrpVal	289
Db	757	GGCTACTACCTGATCCAGATTCATACCCGGCGGATTGATGTTATTTCTTCGGTTC	816
OY	290	SerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrVal	309
Db	817	TCCTTTGGCTCACCGTACCGCTAGTCCAGCTCGCGCTCGCGCTCACACCGTG	876
OY	310	LeuThrMetThrThnGlnSerSerGlySerArgAlaSerLeuProLysValSerTrpVal	329
Db	877	CTCAGCATGACCACTCATGTCCAGTACCACACGCCGCTGCCAAAATATCTACGTC	936
OY	330	LysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIleLeuLeuGlu	349
Db	937	AAGATATTCGACGCTACCTGGGCACATGTTTCGTAAATGTTTACCGGCTCTCGAG	996
OY	350	TyrAlaIleAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLeuArgArg	368
Db	997	TACGCGCGGTAGATATCTCGCAAGAGA-----ATCACCATGAGAAACC	104
OY	369	-----ArgGlnArgArgGlnArgLeuGluGluAspIle	379
Db	1045	CGCTGTCAGCAGCTGGCAAACTTGCAGAGCAACACAGCAGAGATG-CGCCCGGCTTC	110
OY	380	IleGlnGluSerArg-----PheTrpPheArgGlyTrpGlyLeuGlnHisCysLeu	396
Db	1104	TTCCACACGAGCCAGACTCTGAGACCCCTTGCTAGCCAGTCTCTGAGATATC---CATTTGT---	115
OY	397	GlnAlaArgAspGlyGlyPheMetGluGly-----SerGlyIleTrpSerPro	412
Db	1158	-----CAAGACGGTGGTTCCTGTCGTCGGTGGTTCCTGCTGGCGGTGCATC-----	120
OY	413	GlnProProAlaProLeuLeuArgGluGluGluThrThrArgLysLeuTrp	429
Db	1203	-----CCAAAGACAACCGAGGAGACACACCAACCGCATTTTAC	1241
RESULT 14			
AAD22070 standard; DNA; 1614 BP.			
XX	AAD22070;		
AC	AAD22070;		
XX	12-FEB-2002	(first entry)	
DT			
XX			
DE		Dermacentor variabilis clone 8 GABA-gated chloride channel DNA.	
XX			
KW		Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;	
KW		GABA-gated chloride channel; recombinant expression; domestic animal; ds.	
XX			
OS		Dermacentor variabilis.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..1614	
FT		/product="GABA-gated chloride channel protein".	
XX			
PN	W0200174884-A1.		

XX	11-OCT-2001.
XX	
PD	
XX	28-MAR-2001; 2001MO-US09955.
XX	
FE	
XX	31-MAR-2000; 2000US-193791P.
PR	
XX	
XX	(MERI ) MERCK & CO INC.
PA	
XX	
PI	Zheng Y, Cully D, Ludmerer S;
XX	
XX	WPI: 2002-010778/01.
DR	
XX	P-PSDB; AAEI3312.
PT	New polypeptide useful for preventing or treating tick infestation, in
PT	humans, dogs, cattle, horses, deer, or other wild or domesticated
PT	animals, comprises the Dermacentor variabilis gamma-aminobutyric acid
PT	(GABA)-gated chloride channel -
XX	
XX	
PS	Claim 8; Fig 3; 59pp; English.
XX	
CC	The invention relates to gamma-aminobutyric acid (GABA)-gated chloride
CC	channels and their corresponding nucleic acid molecules. GABA-gated
CC	chloride channel proteins and DNA's are useful for preventing and
CC	treating tick infestation, particularly in humans, dogs, cattle, horses,
CC	deer, or other wild or domesticated animals. The nucleic acids are useful
CC	as hybridisation probes or Polymerase Chain Reaction primers for
CC	identifying the presence of Dermacentor variabilis GABA-gated chloride
CC	channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic
CC	acids are also useful for the recombinant expression of D. variabilis
CC	GABA-gated chloride channel proteins. GABA-gated chloride channel
CC	proteins exert toxic effects on other ticks or related parasites such as
CC	miles. The present sequence is Dermacentor variabilis clone 8 GABA-gated
CC	chloride channel DNA.
XX	
XX	
SO	Sequence 1614 BP; 369 A; 475 C; 422 G; 348 T; 0 other:
	Alignment Scores:
	Pred. No.: 4,36e-75 Length: 1614
	Score: 737.00 Matches: 164
	Percent Similarity: 58.78% Conservative: 67
	Best Local Similarity: 41.73% Mismatches: 114
	Query Match: 32.73% Indels: 49
	DB: 24 Gaps: 11
	US-10-075-846-4 (1-431) x AAD22070 (1-1614)
OY	55 ThSerGclgYrTrsApIarIglleArpProAsenPhelySGlyProProvaIAsnValThr 74
DB	154 ACACGCTGGGACACACGAGAGGGGAGGCCAAATVATGGCGGCTCCAGTGAAGTGGC 213
OY	75 CysAsnIlePheIleAsnSerPheSerSerValTrhTrsThrMetAspTYrArgVal 94
DB	214 GTCACTATGCAATVTCAGCATTAAGTACAGTCTCTGAAGTACAATAGTGACTTACTCT 273
OY	95 AsnValPheIleuArgIleIntrTpAsnAspProAtgIleuSerTYrArgGIuTYrProAsp 114
DB	274 GACTTCTATTTCGGGCAATCGTGGGGCGAGCAGCACTCTGTTCCAAAAAAGCCAGAC 333
OY	115 ---AspSerIleuAspIleuAspProSerMetIleuAspSerIleTPlyrProAspIleuPhe 133
DB	334 CTGAGAGCACTGACTGTGGCCGCTGAAGTGGCCGAGAGAGATCTGGTACCCGACACCTTC 393
OY	134 PheIleAsnGIuLSyGIyAlaAsnPhenIleSGluValTrhTrsAspAsnTyLSyLeuIleuArg 153
DB	394 TTGCGCAACAGAGAAGGCGCTTACTTTCATGTGGGGCACACAGCCCAACACTTCTCTCGGC 455
OY	154 IlePheIleuAsnGIyAsnValIleuTYrSerIleatrgIleuIleuIleuSerCysIleu 173
DB	454 ATCGGCTCCGAGAGAGAGGTTTTCGCACTATTTCACATGACGAGGACTGCGACGTGCCCA 513
OY	174 MetAspIleuTyLSyAsnPhenIleAspIleGIIntrhTrsCysTrhMetGIInleuIleuSer 193

XX WO200174884-A1.  
 PN 11-OCT-2001.  
 PD 28-MAR-2001; 2001MO-USO9955.  
 PE 31-MAR-2000; 2000US-193791P.  
 PR (MERI) MERCK & CO INC.  
 PA Zheng Y, Cully D, Ludmerer S;  
 PI WPI: 2002-010778/01.  
 DR P-PSDB: AAE13313.  
 XX New polypeptide useful for preventing or treating tick infestation, in  
 PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
 PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid  
 PT (GABA)-gated chloride channel -  
 PS Claim 8; Fig 3: 59pp; English.  
 XX The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
 CC channels and their corresponding nucleic acid molecules. GABA-gated  
 CC chloride channel proteins and DNA's are useful for preventing and  
 CC treating tick infestation, particularly in humans, dogs, cattle, horses,  
 CC deer, or other wild or domesticated animals. The nucleic acids are useful  
 CC as hybridisation probes or Polymerase Chain Reaction primers for  
 CC identifying the presence of Dermacentor variabilis GABA-gated chloride  
 CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
 CC acids are also useful for the recombinant expression of D. variabilis  
 CC GABA-gated chloride channel proteins. GABA-gated chloride channel  
 CC proteins exert toxic effects on other ticks or related parasites such as  
 CC mites. The present sequence is Dermacentor variabilis clone 9 GABA-gated  
 CC chloride channel DNA.  
 XX Sequence 1614 BP: 371 A; 474 C; 421 G; 348 T; 0 other:  
 SO  
 Alignment Scores:  
 Pred. No.: 1,45e-74 Length: 1614  
 Score: 732.50 Matches: 161  
 Percent Similarity: 57.83% Conservative: 68  
 Best Local Similarity: 40.66% Mismatches: 113  
 Query Match: 32.53% Indels: 55  
 DB: 24 Gaps: 11  
 US-10-075-846-4 (1-431) x AAD22071 (1-1614)  
 QY 55 ThSerGlyTyrAspAlaArgIleArgProAsnPhelYsglyProProValAsnValThr 74  
 DB 154 ACAGGTGGTACGACGAGGAGGTGAGCCAAATATGCGCGCTCCAGAGGAAGTTGGC 213  
 QY 75 CysAsnIlePhelIleAsnSerPheSerValThrIleThrMetAspTyrArgVal 94  
 DB 214 GTCAGTATGAGATATACCAATTAACAGTCTCTCAAGTACAAAGACTTACTTCT 273  
 QY 95 AsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSerTyrArgGlnTyrProAsp 114  
 DB 274 GACTTCTATTTCGCGCAATGCTGGCGGAGGAGGACTCTCTGTCAGAAAAGCCCGAC 333  
 QY 115 ---AspSerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPhe 133  
 DB 334 CTCGAGAGCATGACTGGCGCGCTGAAGTGGCCGAGAGATGTGGTACCGACACCTTC 393  
 QY 134 PheIleAsnGlnLysGlyAlaAsnPhelIleGluValThrThrAspAsnLysLeuLeuArg 153  
 DB 394 TTGCGCACAGAGAGAGCGCTACTTTCATGCGCGCCACCAACACTTTCCTCCGC 453  
 QY 154 IlePheLysAsnGlnLysAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeu 173  
 DB 454 ATCGGCTCCGAGAGAGAGGTTTTCGCAATATTCGACTGACGGTACGCCAGCTGCCCA 513

QY 174 MetAspLeuLysAsnPhelProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSer 193  
 DB 514 ATGATCTCAGATACTCCGATGACAGACAGCGCATTAATGAGATAGAA----- 567  
 QY 194 SerIleLeuCysSerProLeuProSerLeuSerLeuSerValIleGlyThrMetLysAsp 213  
 DB 568 -----AGCTTTGGTTATATACCAAGAAAGAC 591  
 QY 214 LeuValPheGlnTrpLeuGlnAlaAspAlaProAlaValGlnValIleGlnGlyLeuThrLeu 233  
 DB 592 ATCCGCTACCGGTGGTGGACGGTGCACGCTCCGCTCCGCAAGAGAGTGAAGTTG 651  
 QY 234 ProGlnPhe-----IleLeuArgAspIleLysAspLeuGlnLysCysThrLys 249  
 DB 652 CCGCAGTTCAAGGCTCGGTGCACGTCACAAAGCCAAAGAGGTGGCCCTAACGACAGA 711  
 QY 250 HistyrAsnThrIleLysPheThrCysIleGluValLysPheIleLeuGlnMet 269  
 DB 712 AACTACATCC-----CCCTGGTATGT--GAATACGGTTTC-----GCCCGCTCCATG 756  
 QY 270 GLYTyrTrpLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTrpVal 289  
 DB 757 GGCTACTACCTGATCCAGATCTACATCCCGCGGATGATCGTGTATTTCTCGGGTC 816  
 QY 290 SerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrVal 309  
 DB 817 TCCTTTGGGTCCACCGTAAACGATGACAGTCCGCTCCGCGCTCCGCTCCACACCGTC 876  
 QY 310 LeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyrVal 329  
 DB 877 CTCACGATGACCAACATCATGTCAGTACCAACGACGCGTCCCAAAATATCTTACGTC 936  
 QY 330 LysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaAlaLeuLeuGlu 349  
 DB 937 AAGAGTATCAGCATGCTACTGCGCACATGTTTGTATGTGTTTACCGCGCTCTCGAG 956  
 QY 350 TyrAlaAlaIleAsnPhelValSerArgGlnHisLysGlnPheIleArgLeuArgArg--- 368  
 DB 997 TACGCGCGGTAGATATCTCGGCAAGAGA-----ATCACCATGAGGAAGAAC 1044  
 QY 369 -----ArgGlnArgArgGlnArgLeuGlnGluAspIle 379  
 DB 1045 CGCTGTACAGAGCTGCGAAACATTGCAGACCAACAGGAGAGATG-CGCCGACGCTTC 1103  
 QY 380 IleGlnGlnSerArgPheTyrPheArgGlyTyrGlyLeuGly----- 393  
 DB 1104 TTCCACAGACCAAGCTC-----TGAGCCCTTCTACCAAGCTCTCAAGTATC 1151  
 QY 394 HisCysLeuGlnAlaArgAspGlyGlyProMetGlnLysSerGlyIleTyrSerProGln 413  
 DB 1152 CATTTGTCACACGGTGGCTTCGTCAAGTTTGTCCGTGCGT----- 1196  
 QY 414 ProProAlaProLeuLeuArgGlnGlyGlyIleThrThrArgLysLeuTyr 429  
 DB 1197 ---GCAATCCCAAGGACACACAGGAGGAAAGCACCAACCGATTTTAC 1241

Search completed: June 30, 2003, 23:13:41  
 Job time : 383.186 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - nucleic search, using frame\_p2n model

Run on: June 30, 2003, 23:02:29 ; Search time 91.5995 Seconds  
(without alignments)  
1442.995 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252

Sequence: 1 MTLVLPATISFLMLTLPCO.....POPPAPLRGCTTRKLYVD 431

Scoring table: BLOSUM62  
Xgapop 10.0, Ygapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n\_model -DEV=xlp  
-O/cgna2\_1/USPTO.spool/US10075846/runat\_25062003\_163649\_5167/app\_query.fasta.1.782  
-DB=Issued\_Patents\_NA -OFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846.rcgn.1.1.84 -runat\_25062003\_163649\_5167 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NBS\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued\_Patents\_NA:

- 1: /cgna2\_6/ptodata/1/ina/5A.COMB.seq:\*
- 2: /cgna2\_6/ptodata/1/ina/5B.COMB.seq:\*
- 3: /cgna2\_6/ptodata/1/ina/6A.COMB.seq:\*
- 4: /cgna2\_6/ptodata/1/ina/6B.COMB.seq:\*
- 5: /cgna2\_6/ptodata/1/ina/6C.COMB.seq:\*
- 6: /cgna2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	750	33.3	4621	4	US-09-592-891A-13
2	749	33.3	1609	4	US-09-592-891A-11
3	731.5	32.5	1368	4	US-09-130-339-1
4	731.5	32.5	3958	5	US-08-435-933-5
5	731.5	32.5	3958	5	PCT-US96-06035-5
6	728.5	32.3	1640	4	US-09-592-891A-12
7	728.5	32.3	1844	4	US-09-002-361-1
8	726	32.2	2066	3	US-08-072-064-7
9	724.5	32.2	1519	4	US-09-002-361-4
10	724	32.1	1657	1	US-08-554-659-1
11	724	32.1	1657	1	US-08-554-659-3
12	720	32.0	2066	3	US-08-072-064-2

13	720	32.0	2066	3	US-08-072-064-3	Sequence 3, Appl1
14	718	31.9	2066	3	US-08-072-064-5	Sequence 5, Appl1
15	717	31.8	1491	1	US-08-137-614A-3	Sequence 1, Appl1
16	717	31.8	1491	2	US-08-768-301-1	Sequence 1, Appl1
17	712	31.6	1866	1	US-08-417-330A-19	Sequence 19, Appl1
18	687.5	30.5	1970	3	US-08-072-064-9	Sequence 9, Appl1
19	687.5	30.5	1970	5	PCT-US92-08558-3	Sequence 3, Appl1
20	687.5	30.5	2059	5	PCT-US92-08558-2	Sequence 2, Appl1
21	683.5	30.4	1917	4	US-09-627-6508-6	Sequence 6, Appl1
22	683.5	30.4	1917	4	US-09-436-063C-6	Sequence 6, Appl1
23	680.5	30.2	1297	4	US-09-130-339-3	Sequence 3, Appl1
24	680.5	30.2	1555	4	US-08-809-802-11	Sequence 11, Appl1
25	673	29.9	1479	2	US-08-249-112-2	Sequence 2, Appl1
26	673	29.9	1479	5	PCT-US95-06556-2	Sequence 2, Appl1
27	670.5	29.8	1542	1	US-08-249-112-1	Sequence 1, Appl1
28	670.5	29.8	1542	5	PCT-US95-06556-1	Sequence 1, Appl1
29	663	29.4	2601	4	US-09-627-6508-10	Sequence 10, Appl1
30	663	29.4	2601	4	US-09-436-063C-10	Sequence 10, Appl1
31	657	29.2	2544	4	US-09-627-6508-4	Sequence 4, Appl1
32	657	29.2	2544	4	US-09-436-063C-4	Sequence 4, Appl1
33	656	29.1	2508	4	US-09-627-6508-8	Sequence 8, Appl1
34	656	29.1	2508	4	US-09-436-063C-8	Sequence 8, Appl1
35	645	28.6	1732	1	US-08-417-330A-17	Sequence 17, Appl1
36	628	27.9	2635	2	US-08-768-301-3	Sequence 3, Appl1
37	625.5	27.8	1650	1	US-08-459-100A-1	Sequence 1, Appl1
38	625.5	27.8	1650	5	PCT-US94-09589-1	Sequence 1, Appl1
39	619.5	27.5	1707	4	US-08-809-802-7	Sequence 7, Appl1
40	618	27.4	2191	1	US-08-417-330A-11	Sequence 11, Appl1
41	610.5	27.1	1638	1	US-08-417-330A-13	Sequence 13, Appl1
42	603.5	26.8	1652	4	US-09-627-6508-2	Sequence 2, Appl1
43	603.5	26.8	1652	4	US-09-436-063C-2	Sequence 2, Appl1
44	599	26.6	2310	1	US-08-417-330A-15	Sequence 15, Appl1
45	529.5	23.5	669	4	US-09-002-361-7	Sequence 7, Appl1

## ALIGNMENTS

RESULT 1  
US-09-592-891A-13  
Sequence 13, Application US/09592891A  
Patent No. 6329174  
GENERAL INFORMATION:  
APPLICANT: Xiao-Zhou Michelle Wang  
APPLICANT: Xavier Georges Sarda  
APPLICANT: Michael David Tomalski  
APPLICANT: Vincent Paul Mary Wingate  
TITLE OF INVENTION: Heliothis Glutamate Receptor  
FILE REFERENCE: A32815 072667.0118  
CURRENT APPLICATION NUMBER: US/09/592.891A  
CURRENT FILING DATE: 2001-01-22  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 13  
LENGTH: 4621  
TYPE: DNA  
ORGANISM: Heliothis virescens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (764)...(764)  
OTHER INFORMATION: n is a, t, g, or c  
US-09-592-891A-13

## Alignment Scores:

Pred. No.: 1.77e-85 Length: 4621  
Score: 750.00 Matches: 167  
Percent Similarity: 58.42% Conservative: 62  
Best Local Similarity: 42.60% Mismatches: 139  
Query Match: 33.30% Indels: 25  
DB: 4 Gaps: 7

US-10-075-846-4 (1-431) x US-09-592-891A-13 (1-4621)

OY 31 GluVallysserGlyThrlyGlySerClnPrometSerProserAspPheLeuAspLys 50

Db	207	GAATGCATGAAACGGGAGAGATCAACTTTCGAGAACGAGAACGAGATCCCTGGATCAG	266
Qy	51	LeuMetGlyArgThrSerGlyTyrAspAlaArgLeuPro-----AsnPheLys	67
Db	267	ATCCGAGGC---CCCGGAGGTACGACGCCAGATGACACCCCGGGATGTCACGCGCACT	323
Qy	68	GlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerSerValThrLys	87
Db	324	GATGGGCCACGGCTGTGAGCCTCAATATATTTGTCCGAGATGTATCAAGAATCGATGAC	383
Qy	88	ThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnThrPasnProArgLeu	107
Db	384	GTCACAAATGAGATCTCCGTCACAGTTAAAGCTTGGGAGCAATGTTAGATGAACGCTC	443
Qy	108	SerTyrArgGluTyrProAspSerSerLeuAspLeuAspProSerMetLeuAspSerIle	127
Db	444	AAATTCATATATCTTGAGAGTCGCCCTCAATACCTGACACTGACCCGACCAACCAAGAGTC	503
Qy	128	TrpLysProAspLeuPhePheIleAsnGlnLysGlyAlaAsnPheHisGlnValThrThr	147
Db	504	TGGAGTCCGTGATCTATTCTTCTCCAAACGAGAAAGAAAGTCATTTCCAAACATCATCAG	563
Qy	148	AspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThr	167
Db	564	CCGAAAGCTGTACATCCGGATTTCTCCCAACGGCCAAAGTCCTGTACAGCATCCCAATCTCC	623
Qy	168	LeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThr	187
Db	624	TTGAGCGCTCTGCTGCCATCGAACCCTCAAGTTGTACCCCTGGATTAAGCAGACCTGCTCG	683
Qy	188	MetGlnLeuGluSerSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerVal	207
Db	684	CTCAGAGATG-----GCTAGATTAT 701	
Qy	208	GlyTyrThrMetLysAspLeuValPheGluTyrPheGlnAspAlaProAlaValGlnVal	227
Db	702	GGTGGACACACAGACGACTTACTGTCTTATGAGAAAGAAAGCCAGCCG---GTGCAAGTG	758
Qy	228	AlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluAspLeu---GlyCys	246
Db	759	GTGAAATACCTTACACTGCGCTCGTTACACGCG---GAGAAATCTTCACACTGACTAC	812
Qy	247	CysThrLysHisTyrAspThrCysPheThrCysIleGluValLysPheHisLeuGln	266
Db	813	TGCACACTAGACCTAATACCGGTGAATACAGTGTGCTGAAGGTAAACTGTGCTTCCAA	872
Qy	267	ArgGlnMetGlyTyrTyrIleuIleGlnMetTyrIleProSerLeuLeuIleValIleLeu	286
Db	873	CGCGAGTTCAGTTACTACTGATCCAGATCTPACATTCGTCGTCATGCGTCAATCGTC	932
Qy	287	SerTyrValSerPheThrPheLysMetAspAlaAlaProAlaArgValGlyLeuGlyIle	306
Db	933	TCTTGCGGTGCTTGTGCTGACCAAGGAGCTGTGCGCGAGGGTCTTCACATGAGAGTG	992
Qy	307	ThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVal	326
Db	993	ACGACTTACTTACATGACGCGACGACGATGTCAGCATCAACGCGTCCCTACACCGGTC	105
Qy	327	SerTyrValLysAlaIleAspIleThrPheAlaValCysLeuLeuPheValPheAlaAla	346
Db	1053	TCTCTACAGAAAGCCATTGACTGTGAGCTGTGATGTCTCATTCATTCGATTGCGGAGCG	1111
Qy	347	LeuLeuGluTyrTyrAlaAlaIleAsnPheValSerArg-----GlnHisLysGlnPheIle	364
Db	1113	CTACTAGAGTTGCGCCTGCTCACTATGCGCTTCGCTTGACATGCAACCGAGAAACATG	1177
Qy	365	ArgLeuAsnArgArgGlnArgArgGlnArgLeuGlnGluGlnAlaAspIleIleGlnGlnSerArg	384
Db	1173	AAGAAAGGACGGGAGATGAGACGACGACATGAGATGC-TGCCCTCAGATCTCTAGA	1233
Qy	385	PheTyrPheAsnArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaAlaArgAspGlyGlyProMet	404

```

Db      1222 CACAGATGCAACACACCCTTTGGCATGAACACCTTGCTGGCGGCGCGGTGTCGATCC 1229
Oy      405 GlucylSerGlyIleTyrSerProGlnProAla 416
          ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      1292 CAGATGCGGCGAGTGCAGATCCACATCCACCCCGCC 1327
RESULT 2
US-09-592-891A-11
; Sequence 11, Application US/09592891A
; Patent No. 6329174
; GENERAL INFORMATION:
; APPLICANT: Xiao-zhou Michelle Wang
; APPLICANT: Xavier Georges Sarda
; APPLICANT: Michael David Tomaski
; APPLICANT: Vincent Paul Mary Wingate
; TITLE OF INVENTION: Helicobis Glutamate Receptor
; FILE REFERENCE: A32815 072667.0118
; CURRENT APPLICATION NUMBER: US/09/592.891A
; CURRENT FILING DATE: 2001-01-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
;
; LENGTH: 1609
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PCR amplified fragment of Helicobis DNA cloned
; US-09-592-891A-11
; (invitrogen)

```

Alignment Scores:		
Pred. No.:	3,71e-86	Length: 1609
Score:	748.00	Matches: 158
Percent Similarity:	61.56%	Conservative: 55
Best Local Similarity:	45.66%	Mismatches: 109
Query Match:	33.26%	Indels: 24
DB:	4	Gaps: 7
US-10-075-846-4 (1-431) x US-09-592-891A-11 (1-1609)		

OY	3	GlValLysSerGlyThrThrGlySerGlnPrometSerProSerAspPheLeuAspLys	50
Db	211	GAATGCATGACACGGTGGGAAGATCACTTCGAGAAAGAGAACGACAGATCTGGATCAG	270
OY	51	LeuMetGlyAthrThrSerGlyTyrAspAlaArgIleArgPro-----AsnPheLys	67
Db	271	ATCTGGGGC---CCCGGAGAGTACGACGCCAGATACAGCCCTCGGGATCAACGCCACT	327
OY	68	GlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerSerValThrLys	87
Db	328	GATGGCCACCGCTAGTGACCGCTCAATATATTTGTCGGAATATATCAACAGATGATGAC	387
OY	88	ThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnIntTyrAsnAspProArgLeu	107
Db	388	GTCACATGGAAATACACCGACACATTAATACGTTTCGGGAACAATGGTATGATGACACGCTC	447
OY	108	SerTyrArgGluTyrThrProAspSerSerLeuAspLeuAspProSerMetLeuAspSerIle	127
Db	448	AAATTCATTAATCTTGGAGCTCGCTCAATACCTGACACTGACAGGAAGCCACACAGACTC	507
OY	128	TrpLysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGluValThrThr	147
Db	508	TGGATGCCGTGATCTATTCTTCTTCACAGGAAGAAAGCATTTCCACACATCATCATG	567
OY	148	AspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThr	167
Db	568	CCGAACTGTACATCCGAATCTTCCCAACGAGCAACGTGCTACAGCATCCGATATCC	627
OY	168	LeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThr	187
Db	628	CTGACGCTCTGGTCCCATGAAACCTCAAGATTGTACCCCTGGAGAACACACCTGGCTCG	687
OY	188	MetGlnLeuLysSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerVal	207



```
Db      688 CTCAGCATG-----GCTAGTAT 705
Qy      208 GATYrThmetLysaspLeuValPhegluTrpLeuGluaspAlaProAlaValIglInval 227
Db      706 GGTGGACCCAGACAGACTAGTGTCTTATGGAGAGAGCGACCG--GTGCAGGTG 762
Qy      228 AlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysaspLeu---GlyCys 246
Db      763 GTGAAAACCTACACTGCGCTCGCTCAGCGTG-----GAGAAGTTCTCCTACTGACTAC 816
Qy      247 CysThLysHisTrpAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGlu 266
Db      817 TGCACAGTAAGACTATACCGGTGATACAGTGTGCTGAGAGTAGACCTGCTCTTCAA 876
Qy      267 ArgGlnMetGlyTrpTrpLeuIleGlnMetTrpIleProSerLeuLeuIleValIleLeu 286
Db      877 CGCGAGTTCAAGTACACCTGATCCAGATCTACATTCCGTGCTGATGCTGATCCTG 936
Qy      287 SerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIle 306
Db      937 TCTGTGGTGTCTTCTGCGTGGACAGAGAGCTGTGCTGCGAGGTTCTACTAGAGTG 996
Qy      307 ThrThValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysVal 326
Db      997 ACGACTTACTACTACATGCGGACCCAGATCGTCAGAGCATCAACCGCTCCCTACACCGGTG 1056
Qy      327 SerTrpValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIle 346
Db      1057 TCTTACACGAAACCCATTGATGTCTGACTGGGTATGTCTCACTTCGATTCGGAGCG 1116
Qy      347 LeuLeuGluTrpAlaIleAsnPheValSerArg-----GlnHisLysGluPheIle 364
Db      1117 CTACTAGAGTTTGGCTGCTCACTATGCTGTGCTGCTGACATGCACGACGAGAACATG 1176
Qy      365 ArgLeuArgArgArgGln 370
Db      1177 AAGAAAGCGAGCGGAG 1194

RESULT 3
US-09-130-339-1
; Sequence 1, Application US/09130339
; Patent No. 638701
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Paresse, Philip S.
; APPLICANT: Warmke, Jeffrey W.
; APPLICANT: Eiter, Adrian
; APPLICANT: Cohen, Charles J.
; APPLICANT: Brochu, Richard M.
; TITLE OF INVENTION: DNA MOLECULES ENCODING CTENOCEPHALIDES
; TITLE OF INVENTION: FELIS GLUTAMATE GATED CHLORIDE CHANNELS
; FILE REFERENCE: 20029
; CURRENT APPLICATION NUMBER: US/09/130,339
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1368
; TYPE: DNA
; ORGANISM: ctenocephalides felis
US-09-130-339-1

Alignment Scores:
Pred. No.:      5,03e-84      Length:      1368
Score:          731.50      Matches:      157
Percent Similarity: 59.78%      Conservative: 60
Best Local Similarity: 43.25%      Mismatches:  111
Query Match:    32.48%      Indels:      35
DB:              4          Gaps:          9

US-10-075-846-4 (1-431) x US-09-130-339-1 (1-1368)
```

```
Qy      24 ArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnPrometSer 43
Db      79 CGCTAGAGAGCGGCAAGAAATTTTCAGGCGCAAGAAAG----- 120
Qy      44 ProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTrpAlaArgIleArg 63
Db      121 -----CAAGTTGTGACCAAAATTTTAAAGC--CCAGGCCATTACATGCGCAAAATAGG 171
Qy      64 ProAsn-----PheLysGlyProProValAsnValThrcysAsnIlePhe 78
Db      172 CTTTCTGGAGCTCAATGAGACTGAGACGCGTCCACCGTG--GTACAGTCAACATCTAT 228.
Qy      79 IleAsnSerPheSerSerValThrLysThrThrMetAspTrpArgValAsnValPheLeu 98
Db      229 CTGAGATTCATACCGCAAAATAGATGACTACAAATGGAATAGATGTCGACGTTAACTTC 288
Qy      99 ArgGlnGlnTrpAsnAspProAlaGluSerTrpArgGluTrpProAspSerLeuAsp 118
Db      289 AGGGAACAAATGCGCAGATGAGAGTTGAAATTTTAAAGACTTTTGAGAGTCTTTAAATAC 348
Qy      119 LeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAlaAsnGluLys 138
Db      349 TTAACTACACCGACAGCACTGCTGATGAGATGCCGATTTGTCTTGGAAATGAAG 408
Qy      139 GlyAlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGly 158
Db      409 GAGGGCCACTTTCACAACATCATCATGCCGAGCTGCTACATTCGTATTTTCTTACGGT 468
Qy      159 AsnValLeuTrpSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsn 178
Db      469 TCCGTACTATACAGCATCAGATATGCTTACCTTATGCGGTCTTGAATTCGAACATG 528
Qy      179 PheProMetAspIleGlnThrCysThrMetGlnLeuLeuSerSerIleLeuCysSer 198
Db      529 TATCCGCTCATAGCGCAGAGTGTCTCTCCGATG----- 564
Qy      199 ProLeuProSerLeuSerLeuSerValGlyTrpThrMetLysAspLeuValPheGluTrp 218
Db      565 -----GCCAGTTATGCTTGGACCAACAAATCATTCGTGTTTCTG 606
Qy      219 LeuGluAspAlaProAlaValIglInvalIleGluGlyLeuThrLeuProGlnPheIleLeu 238
Db      607 AAGGAAGTGACCG--GTGCAGGTTGTCAAAGAACTTACATCTGCCAGGTTTACCTG 663
Qy      239 ArgAspGluLysaspLeu---GlyCysCysThrLysHisTrpAsnThrGlyLysPheThr 257
Db      664 -----GAGAAGTTCTTGACCGATTTATTAACAGCAAAACCAATACCGTGAATACAGT 717
Qy      258 CysIleGluValLysPheHisLeuGluArgGlnMetGlyTrpTrpLeuIleGlnMetTrp 277
Db      718 TGCCTGAGAGTGCACCGTCTCTTAAAGAGAGTTCGTAACCTGATCCGATCCAGATCTAC 777
Qy      278 IleProSerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAla 297
Db      778 ATTCTTGTTCATGTTGATGATCGTTTCTCGGTGTCGTTGTTGTTGAGCGAGCG 837
Qy      298 AlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrGlnSerSer 317
Db      838 GTTCCGCGCAGAGATATCAGGTGGGTGACCATCTCTCCACCATGCCACCGACGCTCG 897
Qy      318 GlySerArgAlaSerLeuProLysValSerTrpValLysAlaIleAspIleTrpMetAla 337
Db      898 GGCATTAAGCCCTCCCGCCGCGCTCTACACAAAGCCATCGACGTGAGACCGGA 957
Qy      338 ValCysLeuLeuPheValPheAlaIleLeuGluTrpAlaIleAsnPheValSer 357
Db      958 GTCTGCTTACAGTTCGTTCTTGGGGCTTCTCGAAATTCGCCCTCGCACTACGCTTC 1017
Qy      358 Arg-----GlnHisLysGluPheIleArgLeuArgArgArgGlnArgGlnArgLeu 375
Db      1018 AGATCCGATATGCACAGGGAAGAAACATGAAGAAAGCGCAGGAACTTGAACAAGACGCC 1077
Qy      376 GluGluAsp 378
```

Db 1078 AGCTGAC 1086  
RESULT 4  
US-08-435-933-5  
Sequence 5, Application US/08435933  
Patent No. 5693492  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Arena, Joseph P.  
APPLICANT: Pares, Philip S.  
APPLICANT: Liu, Ken K.  
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: John W. Wallen III  
STREET: 126 East Lincoln Avenue  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/435,933  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Wallen, III John W.  
REGISTRATION NUMBER: 35,403  
REFERENCE/DOCKET NUMBER: 19264  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-4720  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3958 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-435-933-5  
Alignment Scores:  
Pred. No.: 3 27e-83  
Score: 731.50  
Percent Similarity: 57.96%  
Best Local Similarity: 42.30%  
Query Match: 32.48%  
Matches: 162  
Conservative: 60  
Mismatch: 127  
Indels: 35  
Gaps: 8  
US-10-075-846-4 (1-431) x US-08-435-933-5 (1-3958)  
QY 3 ThrLeuValProAlaThrLeuSerPheLeuLeuLeuThrProGlyGlnValLeu 22  
Db 390 ACCCCATTCCCAACATGGGACGACACTATTTCGGCGC-----ATCTTA 437  
QY 23 LeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnProMet 42  
Db 438 TACTTTCGACCGCTGCTGCTTCACTACGCAATATATGCAAGTAAATTCGAGAA 497  
QY 43 SerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLe 62  
Db 498 AAGGAGAAAAAGCTTAGATCAATTTTAGT---CGAGCAATACGACGCCGGAATA 554  
QY 63 ArgPro-----AsnPhelysGlyProProValAsnValThrCysAsnIlePheIle 79  
Db 555 CGACCATCTGGAATTAATGGCAGCATGTGTCGCCCATAGTCAGCAATCAATCTATCTGT 614

QY 80 AsnSerPheSerSerValThrLysThrPheLeuAspTyrArgValAsnValPheLeuArg 99  
Db 615 CGCAGTATTATGACGATAGTATGATTAATAATGAGTAGTACGAGTTACCTCCGT 674  
QY 100 GlnGlnTrpAsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeu 119  
Db 675 GAACGTGACGAGGATGAAGCCCTCAAGTT-CGACGATATCCAGCGCTGCCCTAAAGTATCT 733  
QY 120 AspPro-SerMetLeuAspSerIleTrpLysProAspLeuPheLeuAsnGluLysGln 139  
Db 734 GACCTGACGAGCGAAGCCCGCTGTGATCCCGATCTTTCTTCTCGAAGCAGAGGA 793  
QY 139 ValAsnPhelysGluValThrTrpAspAsnLysLeuLeuArgIlePheLysAsnGlyAs 159  
Db 794 GGAACACTTCCCAACATCATCATGCCCCATGTGATATTCGATCTTCCCAACGATC 853  
QY 159 nValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetLysLeuLysAsnPh 179  
Db 854 TGTGTATATAGTATACATCTGCTGACATTTGGCCCTGACCAATGAACCTAAAGCTGTA 913  
QY 179 eProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerTr 199  
Db 914 TCCCGTGTATGACAGATCTGCTCTACGAGATC----- 947  
QY 199 OleuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLe 219  
Db 948 -----GCCACTATGCTGTGAGCACCAACAGCATGGTCTTCCTGCGAA 991  
QY 219 uGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuAr 239  
Db 992 GGAGGGCGATCC---GTACAGGTGTAAAGAACTTACCTACCTGCTTCACACATG- 1046  
QY 239 gAspGluLysAspLeu---GlyCysCysThrLysHisTyrAsnThrGlyLysPheThrC 258  
Db 1047 ---GAGAAAGTTTCTGACTGATTAAGTAAAGAAACCAACCGGTGATACAGTTCG 1102  
QY 258 sIleGluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrI 278  
Db 1103 CCTCAAAAGTCGATCTACTATTACAGCGAAGATTCATATTAATCAATAATTTATAT 1162  
QY 278 eProSerLeuLeuLeuValIleLeuSerTrpValSerPheTrpIleAsnMetLysAlaI 298  
Db 1163 ACCATGCTGTATGTGCTATGATCATGATGATGATGATGATGATGATGATGATGATGAT 1222  
QY 298 aProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrProGlnSerSerG 318  
Db 1223 ACCGCGGAGCTGTCACGTGGTGTACACACCCCTGTCGATGCCACCCAGAGTCGCG 1282  
QY 318 ySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaVal 338  
Db 1283 CATAAAGCCCTCCCTGCGCCGCTTCTTATACGAAGGCCATCGATGTGTGACAGCGGT 1342  
QY 338 lCysLeuLeuLeuValPheAlaAlaLeuLeuGluTyrAlaAlaIleAsnPhelysArg 358  
Db 1343 GTGTCTGACGTTCGTGTCGGGGCTGCTGACGTTCGCTGTCGTCGTCGTCGTCGTCGTC 1402  
QY 358 g-----GlnHisLysGlnPheIleArgLeuArgArgArgGlnAr 371  
Db 1403 ATCAGGTTCGAATTAAGCTAATCATGCAATAGAGAAATATGAAAGAACGCCCGCATCT 1462  
QY 371 gArgGln 373  
Db 1463 CGACGAC 1469  
RESULT 5  
PCT-US96-06035-5  
Sequence 5, Application PC/TUS9606035  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Arena, Joseph P.  
APPLICANT: Pares, Philip S.  
APPLICANT: Liu, Ken K.  
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE



US-09-592-891A-12

Alignment Scores:

Pred. No.:	1,69e+83	Length:	1640
Score:	728.50	Matches:	156
Percent Similarity:	61.108	Conservative:	56
Best Local Similarity:	44.968	Mismatches:	110
Query Match:	32.358	Indels:	25
DB:	4	Gaps:	8

US-10-075-846-4 (1-431) X US-09-592-891A-12 (1-1640)

QY	31	GIValIySSerGlyThrIySGlySerGlnProMetSerProSerAspPheLeuAspLys	50
Db	154	GAATGATGAACGGTGGGAAGATCAACTTCGAGGAAGAGAAAGCAAGATCTGGATCAG	213
QY	51	LeuMetGlyAraGthrSerGlyTyraSpsAlaArgIleArgPro-----AsnPheLys	67
Db	214	ATCTCTGGG---CCCGGAGGTAGACAGCCAGAGATCAACCCCTGGGGATCAACGGCACT	270
QY	68	Gly----ProProValAsnValThrcysAsnIlePheIleAsnSerPheSerSerValThr	86
Db	271	GGCTATGGCCCAACGTTACTGATCATGTCACACATGTTCTACGGTCCATCCAGCAAAATGAT	330
QY	87	LysThrThrMetCaspTyraArgValAsnValPheLeuArgGlnGlnThrPasnAspProArg	106
Db	331	GATTACAAAATGGAAATACCTCCATCAATTAACGTTTCGGGAAACAATGGTTAGATGAACGG	390
QY	107	LeuSerTyraArgLutIyProAspAspSerIleuAspIleuAspProSerMetLeuAspSer	126
Db	391	CTCAAAATTCATTAATCTTGAGGTGGCCCTCAAAATCTCGACACTGATCGAAGCCACACAGA	450
QY	127	IleTrpLysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheAsnGluValThr	146
Db	451	GTCTGATGCTGATCTTATCTTCTCCAAAGAGAGAGAGAGATTTCCACACATCATC	510
QY	147	ThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTySerIleArgLeu	166
Db	511	ATGCCGAAGCTGATCATCCGGATCTTCCCCAAGCCAAAGCTGCTGTATGAGATCCGAATC	570
QY	167	ThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCys	186
Db	571	TCCCTGACCTCTCGTGGCCCATGAACCTCAAGTTGTACCCCTGGATTAACACACCTGC	630
QY	187	ThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSer	206
Db	631	TGCTCATGGATG-----GCTAAT	648
QY	207	ValGlyTyThrMetLysAspLeuValPheGluTrpLeuGluAspAlaProAlaValGln	226
Db	649	TATGATTGGACCCACACAGCACTTAGTGTTCTTATGAGAGAAAGGAGCCG---GTGCAG	705
QY	227	ValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeu---Gly	245
Db	706	GTGTGTAATAAATCTTACCTGCTGCTGCTACGCTG-----GGAATCTCTACTGCAC	755
QY	246	CysCysThrLysHisIleTyraSerThrGlyLysPheThrCysIleGluValLysPheHisLeu	265
Db	760	TACTGCACAGTAAGACTAATACCGGTGAATACATTCCTCGAAGATGACCTCTCTC	819
QY	266	GluArgGlnMetCylTyTyTrpLeuIleGlnMetCylTrpIleProSerLeuIleValIle	285
Db	820	AAACGGAGATCGATTACTACTGATTCACAGATCTCATCTCCGTGCTGATCTGCTGATC	879
QY	286	LeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyGly	305
Db	880	GTGTCTGGGTGCTTCTTGCTGGCCACGAGAGCTGTGCTCGGAGGGTCTCACTAGA	939
QY	306	IleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLys	325
Db	940	GTGACGACATTACTTACATGGGACCAAGCTGTGACGATCAACGCGTCCCTACACCG	999
QY	326	ValSerTyValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAla	345

Db	1000	GTGTCCTACAGCAAGACCATGATGATCTCGAGACTGGGTTATCTCTACACTTGGTATTCCGA	105
Qy	346	AlAlLeuLeuGluTyrAlaAlaIleAsnDheValSerArg-----GlnHisLysGluDhe	363
Db	1060	GCGCTACTAGAGTTTGGCCCTCGTCAACTATATCGTCTCGCTGTGCAATGCAACCGAGAGAAC	1119
Qy	364	IleArgLeuArgArgArgGln	370
Db	1120	ATGAAGAAGACGCAAGCGGAG	1140

RESULT 7

US-09-002-361-1  
; Sequence 1, Application US/09002361  
Patent No. 6320516

```

: GENERAL INFORMATION:
: APPLICANT: Halling, Blaik
: TITLE OF INVENTION: Lepidopteran GABA-Gated Chloride
: TITLE OF INVENTION: Channels
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:

```

STREET: 997 Lenox Drive, Building 3, Suite 100  
CITY: Lawrenceville  
STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEO for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002.361  
PRINT NUMBER:

```

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29.1.1
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1444 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 104...151
OTHER INFORMATION:
US-09-002-361-1

```

Alignment Scores:	
Pred. No.:	2,08e-83
Score:	78.50
Percent Similarity:	57.77%
Best local Similarity:	38.83%
Query Match:	32.35%
DB:	4
Gaps:	10
Indels:	56
Matches:	160
Conservative:	78
Mismatches:	119
Length:	1844

US-10-075-846-4 (1-431) X US-09-002-361-1 (1-1844)

```

QY      17  LeuProG1GlnValLeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThr 36
          |||||  :::::  ::  ::::|  ||
Db      162  TTGCGTGTACCTCATGTGTCACCAATGCCGCGGAGCGGAGAGAG-----GGGGGATGT 215

```

Oy		37	LysGlySerGI nPrometSerProseraspPheLeuAspLysLeuMetCyl nArgThrSer	56
Dd		216	TT -GGTGAC-----GTCAATATTCCAGCATTTTGGATTGCCTA-----ACTGA	259
Oy		57	GlyTy rAspAlaArgIleArgProAsnPhelsglyProProValAsnValThrcysasn	76
Dd		260	AGCTACGACAAMAAGAGTGAGGCCGAACATA TGAGAGACC GCCCGACTGGATTTGGGAGTCACC	319
Oy		77	IlePheIleAsnSerPheSerSerValThrLy sThrThrMetAspTy rAr yValasnaVal	96
Dd		320	ATGTACGGCTCTCCATCAGCTCCCTTATCTGAAGAATAATGATTCCACCTGGATTTTC	379
Oy		97	PheLeuArgGI nGI nTrpAsnAspProAlaGluSerTy rArgGluTy rProAsp---Asp	115
Dd		380	TACTTCACACAACTTTTGGACACACCCCAGCGCTTGCTTACAAAAAAGAGCGGCTGTGGAG	439
Oy		116	SerLeuAspPheLeuAspProserMetLeuAspSerIleTy rLy sProAspLeuPhePheAla	135
Dd		440	ACTCTGTCGCGTGGCGTCCGAATTTATTAGAAACATATGGGTACCCGACACTTCTTTGTT	499
Oy		136	AsngIuLy sgLy AlaAsnPhenHlsgLuValThrThrAspAsnLy sLeuLeuArgIlePhe	155
Dd		500	AACGAAAACAGCTCTTATTTCACATWAGCTACTACAGACAGAAATTCATACGACATTCA	559
Oy		156	LysAsnGI nLy sAsnValLeuTy rSerIleArgLeuThrLeuIleLeuSerCy sLeuMetAsp	175
Dd		560	CATTCTGGATCTATTACTAGAGAGTATTAACACAGACTATATACCGCTTCTTCGATGGAT	619
Oy		176	LeuLy sAsnPheProMetAspIleGI nThrCy sThrMetGI nLeuGI nU SerSerIle	195
Dd		620	TTGCAGTATTTTCCATGATGACCGCTCAATTATGCAATTTGAAATCGAA-----	667
Oy		196	LeucY sSerProLeuProSerLeuSerLeuSerValGI lTy rThrMetLy sAspleuVal	215
Dd		668	-----AGTTTGGCTACACCAVCGCGGACATCCGA	697
Oy		216	PheGI uTP rPLeuGI nAspAlaP rAlaValIGlnValAlaIGluLy sLeuThrl euProGI n	235
Dd		698	TACAGTGAATGAGGGGCCAACACAGTGGGTGTGTCCAGGAAAGATGTCTTTGCCGCAA	757
Oy		236	PheIleLeuArgAspGI nLy sAspleuGI y-----Cy sCy sThrl yShSty r	251
Dd		758	TTC-----AAGTGTCTGGGCCACCGCACCGGCGCATTCGAGATTTCTTT	802
Oy		252	AsnThrGI ly Ly sPheThrCy sIleGI nValLy sPheHisLeuGI nU ArgIleMetGI Ty r	271
Dd		803	ACGACAGAGAACTACTCTGCTGCGATGTGAATTCATTTGTACGTCGCGATGGGATAC	862
Oy		272	TyrLeuIleGI nMetTy rIleProSerLeuLeuIleValIleLeuSerTy rValSerPhe	291
Dd		863	TATTTAAATTCACATTTTATTTCCGCTCGGCTCAATATGTATTCATCTTGGGTATCATTT	922
Oy		292	TrpIleAsnMetAspAlaIal aP rAlaArgValIGly LeuGI yIleThrThrValLeuThr	311
Dd		923	TGGTTGAATCGAAATGCGACACTCGACAGGATATCACTAGGTGTCAACATGTATTGACG	982
Oy		312	MetThrThrGI nSerSerGI ySerArgAlaserIleProLy sValSerTy rValLy sAla	331
Dd		983	ATGACGACGCGTCATGTCGTCACAGAAATGGCGCTGTCCCAAGATTCATATGTCAAGTCC	104
Oy		332	IleAspIleTy rMetAlaValCy sLeuLeuPheValPheAlaIal aLeuLeuGI uTy rAla	351
Dd		1043	ATCGATGTCTATCTGGGAATGTGTTCGTTCATGCGTCTTGCCCGATTACTAGAAATATGCC	110
Oy		352	AlaIleAsnPhelValSerArgGI nHisLy sGI nPheIleArgLeuArgArgValArg	371
Dd		1103	ACGGTTGGCTATATGGCTTAA-----AGGATACAGATGACGAACA	114
Oy		372	ArgGI nArgLeuGI nU AspIleIleGI nGI nU SerArgPheTy rPheArgIly Ty rGI y	391
Dd		1145	AGATTTCACGTCGTTCCAAAATAATGGCGCGCGAGAAATAATGCAATA-----	119
Oy		392	LeuGI nHisCy sLeuGI nAlaArgAspGI yGI yPrometCyl nU SerGI yIleTy rSer	411

```

Db      1193  -----GlnProProAlaPro 417  ||||| |||
OY      412 Pro-----GlnProProAlaPro 417  |||
      11229 CCGAGGACACACCCCTATCTAGGCGCACACACCT 1264  ::||| |||

RESULT 8
US-08-072-064-7
; Sequence 7, Application US/08072064
; Patent No. 6008046
; GENERAL INFORMATION:
; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,064
; FILING DATE: 19930602
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPND-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2066 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
US-08-072-064-7

Alignment Scores:
Pred. No.: 5.33e-83 Length: 2066
Score: 726.00 Matches: 156
Percent Similarity: 56.94% Conservative: 86
Best Local Similarity: 36.71% Mismatches: 143
Query Match: 32.24% Indels: 40
DB: 3 Gaps: 9

US-10-075-846-4 (1-431) x US-08-072-064-7 (1-2066)
OY      5 ValProAlaThrLeuSerPheLeuLeuLeuTPrPhtLeuProGlyGlnValLeuLeuArg 24
      137 CTGCCCGCCGACACA---CCGCTGCTAATACCATCTG-----CTGCCCATCAAC 178
      25 ValAlaLeuAlaLysGluGlu-----ValLysSerGlyThr 36
      179 ATGGCCCTGATTCACACAGAAACGGGCCCAACAGCATCATACAGTGCACACGGCGAGCT 238
OY      37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetGlyLysThrSer 56
      239 GCGGCTGGCAGCATCTGGGTGACGTAACATATCCGCTATTCTGACTCCCTTTAGTGTT 298

```



```
QY 117 LeuAspLeuAspProSerMetLeuAspSerIleThrLysProAspLeuPheAlaAsn 136
    ||| : : : : : ||| ||| ||| |||
Db 319 TTAATGTGGGCGACAAATTCATTAAGAAAGAGGATGCCACGCTTGTAAAT 378
QY 137 GluLysGlyAlaAsnPheHISGLuValThrThrAspAsnLysLeuLeuArgIlePheLys 156
    ||| ||| : : : ||| ||| ||| : : : |||
Db 379 GAAAGCAATCTTATTTCCATATAGCAACACAGCAATGAATTCATCCGTAAACCTAT 438
QY 157 AsnGlyAsnValLeuThrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeu 176
    : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 TCGGCTCTATCTAGAAAGTATACATTCAGATACAGCTCTTGCCCGATGAATTTG 498
QY 177 LysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuLysSerSerIleLeu 196
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 CATACTTCCCGATGATCGACAGTGTGCCACATGAATGA----- 543
QY 197 CysSerProLeuProSerLeuSerLeuSerValGlyThrThrMetLysAspLeuValPhe 216
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 -----AGTTTCGGCTACACCATTCGGCGGACATCAGATAC 576
QY 217 GluThrLeuGluAspAlaProAlaValGlnValAlaGluLysLeuThrLeuProGlnPhe 236
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 AATGGAAGAGAGGCGCAACTGTGTGGGTGTTCCAGCGAAGTGTCCGCGCAGATTG 636
QY 237 IleLeuArgAspGluLysAspLeuGly-----CysCysThrLysHisLysAsn 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 -----AAGGTGCTGGGTATCCGCCAAGAGCTATGAGATCTCCCTTACT 681
QY 253 ThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyThr 272
    ||| ||| : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 682 ACAGGAAATTAATTCACGGTGGCATGGAATCAATTCGTTGCGCTAAGGATATTAC 741
QY 273 LeuIleGlnMetThrIleProSerLeuLeuIleValIleLeuSerThrPheLeuPheTrp 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 742 TTAATCCAATTATATTCCTCTGTTGTTGATTCATCATCATGATATTTGG 801
QY 293 IleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMet 312
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 802 TTGAATCGAAATTCACACAGCTCGAGTGGCCCTCAGTTTCACTGATTTGACAAATG 861
QY 313 ThrThrGlnSerSerGlySerAlaGlnSerLeuProLysValSerThrValLysAlaIle 332
    ||| ||| ||| : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 862 ACAACGCTTATGCTCTACTAACGCGGCGCTGCCAAGATCTCATACGCAATCCATA 921
QY 333 AspIleThrMetAlaValCysLeuLeuPheValPheAlaIleLeuGluIleValAla 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 922 GATGTATATCTGGGACATGTTCTGTCATGTATTTGCTAGTACTAGATACGCGACT 981
QY 353 IleAsnPheValSerArg-----GlnHisLysGluPheIleArgLeuArgArg 368
    : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 982 GTGGATATATGCAAGAAATACAGATGAGAAACAAAGATTTTGCGCCATACAGAA 1041
QY 369 ArgGlnArgArgGlnArgLeuGluAspIle-----IleGlnLysSerArgPheThr 386
    : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1042 ATAGCTTCGAAAGAAATCCCGCTTGACTGCCACCGGTAGCATCCATCACTTTTA 1101
QY 387 PheArgGlyTyrGlyLeuGlnHisCysLeuGlnAlaArg----- 399
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1102 TCAGAGATGGGAACACTTGGCAGATGCCACCGGATGAGACATCGGAGCGGTTCAAA 1161
QY 400 -----AspGlyGlyProMetGlu----- 405
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1162 GTGCATGACCCAAAGCCGATTCCAAAGGGGAGCGAGCTTAGAGAAACATTAATGAGGT 1221
QY 406 -----GlySerGlyIleThrSerProGlnProProAlaProLeuLeuArgGluGlu 423
    : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1222 CGCAGTGGAGCAAGAAAGAAACCAAGCGCCGCCACACATTTTACATCCCGGCAAG 1281
QY 424 ThrThrArgLysLeu 428
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1282 GACATAGCAAACTG 1296
```

```
US-08-554-659-1
: Sequence 1, Application US/08554659
: Patent No. 5767261
: GENERAL INFORMATION:
: APPLICANT: Wingeat, Vincent
: TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Scully, Scott, Murphy & Presser
: STREET: 400 Garden City Plaza
: CITY: Garden City
: STATE: New York
: COUNTRY: USA
: ZIP: 11530-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/554,659
: FILING DATE:
: CLASSIFICATION: 436
: ATTORNEY/AGENT INFORMATION:
: NAME: Digilio, Frank S
: REGISTRATION NUMBER: 31,346
: REFERENCE/DOCKET NUMBER: 9732
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 742-4343
: TELEFAX: (516) 742-4366
: TELEX: 230 901 SANS UR
: INFORMATION FOR SEQ. ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1657 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 115..1581
: US-08-554-659-1
:
: Alignment Scores:
: Pred. No.: 6,54e-83 Length: 1657
: Score: 724.00 Matches: 165
: Percent Similarity: 55.68% Conservative: 80
: Best Local Similarity: 37.50% Mismatches: 115
: Query Match: 32.15% Indels: 81
: DB: 1 Gaps: 12
:
: US-10-075-846-4 (1-431) x US-08-554-659-1 (1-1657)
:
: QY 17 LeuProGlyGlnValLeuLeuArgValAlaLeuAlaLysGluLysSer---Gly 35
: ||| ||| ||| : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 173 TTTCGCTGTTACTCATCTGACCATGCGCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 232
:
: QY 36 ThrLysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThr 55
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 233 ACCTCAATATCTCAGCA-----TTTGG-GATTTCGCTA-----ACT 267
:
: QY 56 SerGlyTyrAspAlaArgIleArgProAsnPheLysGlyProProValAsnValThrCys 75
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 268 GTAGAGTACGACAAAGAGTGAAGCCGAACTATGAGAGGAGCCAGTGTGGAGATC 327
:
: QY 76 AsnIlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsn 95
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
: Db 328 AACATGACGTGCTTCATCAGCTCTTATTCGAGTGAATGATTTTACCCCTGGAT 387
:
: QY 96 ValPheLeuArgGlnGlnThrPasnAspProArgLeuSerTyrArgLysProAsp--- 114
: : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

Db      388  TTCTACTTCAGACAAATTTTGGACAGACCCAGGCTTCTTACAAAAGCAGCGGTGTG 447
Qy      115  Aspsertleuaspleuasprosermetleuaspsertleuplysproaspleuphe 134
Db      448  GAGACTGTGCGCGCGGTGGGAATTTATTAAGAAACATATGGGATCCGACACCTTCTT 507
Qy      135  Alaasnlglyslglaaaspheliasgluvalthrthraspanslysleuarg1le 154
Db      508  GTTAACGAAAAACAGCATATTTCCACATAGCTACTACAGACAGCAATTCATACGAT 567
Qy      155  Pheylasnglyanvalleuylserleargleuthrleuileuileuileuileuileu 174
Db      568  CACATCTCTGATCTATTTACTAGAGATTAAGACTGATACCCGCTTCTTGTCCGAT 627
Qy      175  Aspleuylasnphepmetleasp1leg1nthrcysrhmetgluileuileuileuileu 194
Db      628  GATTTGAGATATTTCCGATGACCGCAATATGCAATATGAAATCGAA----- 678
Qy      195  Ileucysserproleuproserleuileuileuileuileuileuileuileuileu 214
Db      679  -----AGTTTGGCTACACCATCGGAGCATC 705
Qy      215  ValpheglurleugluaspalaProalaValgluValalaugluileuThrleuPro 234
Db      706  CGATACAGATGGAATGAGGGGCCCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 765
Qy      235  Glupheileuargaspglulysaspleugly-----CysCysThrLysHis 250
Db      766  CAATTC-----ANGGTGTGGCGCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 810
Qy      251  TyrasnThrlyslspherhrcysilegluVallysleuileuileuileuileuileu 270
Db      811  CTACGACAGGAACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
Qy      271  TyrThrleuileuileuileuileuileuileuileuileuileuileuileuileu 290
Db      871  TACTATTAAATTCGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 930
Qy      291  PheThrleuasmetaspaalaProalaarValgluileuileuileuileuileuileu 310
Db      931  TTTTGGTTGAATCAAAATGGACACCTGCAAGGCTATCAGTGCACACACTGATTTG 990
Qy      311  ThrmetThrThrGlnserSerGlySerArgAlaSerleuProlysValSerThrVallys 330
Db      991  ACGTGTGACGACGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1050
Qy      331  AlaileaspietpmetAlaValCysleuileuPheAlaAlaileuileuileuileu 350
Db      1051  TCCATTCATCTCTATCTGGCACTTGTTCCTCATGCTGCTGCTGCTGCTGCTGCTG 1110
Qy      351  AlaAlaileasnpheValSerArgGlnHislysluPheleuileuileuileuileu 370
Db      1111  GCCACGGTTGGCTATATGGCGAA-----AGCATACAGATAGGAGAA 1152
Qy      371  ArgArg-----GlnArgLeuGluLysP----- 378
Db      1153  CAAGATTCACACTGCTGTTCAAAAATGCAATATAGATGCTCTCCAGGGTCAAGTCA 1212
Qy      379  ----- 383
Db      1213  ATCCCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1272
Qy      384  ArgPhe-----TyrPheArgLysLysLysLysLysLysLysLysLysLysLys 396
Db      1273  CGGTTCAAAAGTCACGATTCGCAAGGCAATTTCTTAAGCGGTACTTGAAGAAACAT 1332
Qy      397  GlnAlaArgaspGlyGlyPrometGluGlySerGlyLysSerProGlnProProala 416
Db      1333  AATGGGCTGCGGCGCGAGCA-----GGACCTGCTCCACCGGCA 1371

```

RESULT 11  
US-08-554-659-3  
Sequence 3, Application US/08554659

```

Patent No. 5767261
GENERAL INFORMATION:
APPLICANT: Wingate, Vincent
APPLICANT: Wolff, Mark
TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL
TITLE OF INVENTION: AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,659
FILING DATE:
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9732
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1657 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 115..1581
US-08-554-659-3
Alignment Scores:
Pred. No.: 6,54e-83 Length: 1657
Score: 724.00 Matches: 165
Percent Similarity: 55.68% Conservative: 80
Best Local Similarity: 37.50% Mismatches: 115
Query Match: 32.15% Indels: 81
DB: 1 Gaps: 12
US-10-075-846-4 (1-431) x US-08-554-659-3 (1-1657)
Qy      17  LeuProGlyGlnValLeuLeuArgValAlaLeuAlaLysGluGluValLysSer---Gly 35
Db      173  TTGGCTGTACCTCAGTCATGACCATGCGGGGAGCGGAGGAGGAGGAGGAGGAGGAG 232
Qy      36  ThrLysGlySerGlnProMetSerProSerAspPheLeuaspLysLysMetGlyArgThr 55
Db      233  ACGTCATATCTCGCA-----TTTTC-GATTTCGCTA-----AGT 267
Qy      56  SerGlyTyrAspAlaArgLeuArgProasnPheLysGlyProProValasnValThrCys 75
Db      268  GTAGCTACGCAAAAGAGTGAAGGAGCGGAGCTATGAGAGCGGCGAGTGTGGAGTGC 327
Qy      76  AsnIlePheIleasnSerPheSerSerValThrLysThrThrMetAspTyrArgValasn 95
Db      328  AACATGTACGCTCTCATCAGCTCCTTATCTGAAGTGAATAATGATTTACACCTGAT 387
Qy      96  ValPheLeuArgGlnGlnTPAsnAspProArgLeuSerThrArgGluTyrProAsp--- 114
Db      388  TTCTACTTCAGACAAATTTTGGACAGACCCAGGCTTCTTACAAAAGCAGCGGTGTG 447

```



```

QY 115 AspSerLeuAspLeuAspProSerMetLeuAspSerIleTyrPlyProAspLeuPhe 134
Db 448 GAGACTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG
QY 135 AlaAsnGlySGLyAlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIle 154
Db 508 GTTAAAGAAAACAGCATATTTCACATAGCTACACAAAGCAAAATTCATACCAATT 567
QY 155 PheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeuLeuLeuSerLysLeuMet 174
Db 568 CATCATTTCTGGATCTTACTAGAGCTATPAGACTACTACTACCCCTCTTGTCCGATG 627
QY 175 AspLeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSer 194
Db 628 GATTTCAGATATTTCCGATGAGCGCTCATATTATGCAATATTCGAATTCGAA----- 678
QY 195 IleLeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeu 214
Db 679 -----AGTTTGGCTACACCAATGCGGAGACATC 705
QY 215 ValPheGluTyrPheGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuPro 234
Db 706 CGATTACAGTGGATAGGGGCCCAACTAGTGGTGTGCGAGCAAGTGTCTTGGCCG 765
QY 235 GlnPheIleLeuArgAspGluLysAspLeuGly-----CysCysThrLysHis 250
Db 766 CATTTCT-----AAGTGTCTGGCGCCAGCGGCGCCATGGAGATTTCT 810
QY 251 TyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGlnMetGly 270
Db 811 CTTCACACAGCAAACTACTCTCGTGGCATGTGAATTCATTTGATGAGCTCGATGGGA 870
QY 271 TyrTyrLeuIleGlnMetGlyIleProSerLeuLeuIleValIleLeuSerTyrValSer 290
Db 871 TACTATTATTAATTCAGATTATATTCGCTGCGCTAATGTCTATTCCTTGGGATCA 930
QY 291 PheTyrIleAsnMetAspAlaAlaProAlaArgValGlyLysLeuGlyIleThrThrValLeu 310
Db 931 TTTTGGTTGAATGGAATGCGACATGCGAGGCTGAGCACTAGTGTGCACAACTGATTC 990
QY 311 ThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyrValLys 330
Db 991 ACATATACACAGCGCTCATGTCGTCACGAATGCGCGCTCTGCGCAAGATCTCATATGTCAG 1050
QY 331 AlaIleAspIleThrPheAlaValLysLeuLeuPheValPheAlaAlaLeuLeuGluTyr 350
Db 1051 TCCATGATGTCTATCTGGGAACCTGTTCGTCATGCTCTCACCAAGTTTACTAGATAT 1110
QY 351 AlaAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLeuArgArgGln 370
Db 1111 GCCAGGTTGGCTATATGGCGAAA-----AGCATACAGATGAGGAAA 1152
QY 371 ArgArg-----GlnArgLeuGluGluAsp----- 378
Db 1153 CAAGAGTTCACCTGCTCTCAAAAAAATGAATAGTGTCTCCAGGTCAGTGAAGCT 1212
QY 379 -----IleIleGlnGlnSer 383
Db 1213 ATCCCCCAGCAGACACACCCATCTAGGCCACACCATAGCCGATATGAGAGTT 1272
QY 384 ArgPhe-----TyrPheArgGlyTyrGlyLeuGlyHisCysLeu 396
Db 1273 CGGTTCAAGTTCACGATCCGAAGGATATTTCAAAAGCGGTACTTTAGAAAACACTATC 1332
QY 397 GlnAlaArgAspGlyLysProMetGluGlySerGlyIleTyrSerProGlnProAla 416
Db 1333 AATGGGCTCGGGCGCAGCA-----GGACCTGCTCCACCGCA 1371

```

RESULT 12  
 US-08-072-064-2  
 ; Sequence 2, Application US/08072064  
 ; Patent No. 608046  
 ; GENERAL INFORMATION:

```

; APPLICANT: FRENCH-CONSTANT, RICHARD H.
; APPLICANT: JACKSON, MEYER B.
; TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETER G. CARROLL
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/072,064
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 770,881
; FILING DATE: 04-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: OPND-00574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/705-8410
; TELEFAX: 415/397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2066 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; US-08-072-064-2

Alignment Scores:
Pred. No.: 3.16e-82 Length: 2066
Score: 720.00 Matches: 155
Percent Similarity: 56.71% Conservative: 86
Best Local Similarity: 36.47% Mismatches: 144
Query Match: 31.97% Indels: 40
DB: 3 Gaps: 9

US-10-075-846-4 (1-431) x US-08-072-064-2 (1-2066)
QY 5 ValProAlaThrLeuSerPheLeuLeuTyrPheProGlyGlnValLeuArg 24
Db 137 CTGCCCCGACCA---CCGCTGTAACCATCTGG-----CTGGCCATCAAC 178
QY 25 ValAlaLeuAlaLysGluGlu-----ValLysSerGlyThr 36
Db 179 ATGGCCCTGATTCACAGCAAGAGGCGCCACAACGATCCATACAGCGCAGCGGACT 238
QY 37 LysGlySerGlnProMetSerProSerAspPheLeuAsnLysLeuMetGlyArgThrSer 56
Db 239 GCGCGTGGCAGCATGCTGGGTGAGTAACATATCCGCTATTCGACACTCTTAACTGT 298
QY 57 GlyTyrAspAlaArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsn 76
Db 299 AGTTACGACAAAAGAGTAAGCAATTAAGGTGTCCTGCGGAGTTGGCGCTCACCA 358
QY 77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnVal 96
Db 359 ATGTATGTCTCTAGTATCATCTGCTGTTTGGAGGTTCTATGACTTACATTTGATTTT 418
QY 97 PheLeuArgGlnIleThrPheAsnProArgLeuSerTyrArgGluTyrProAsp---Asp 115
Db 419 TACTTCTGCAATTTTGGACGATCTCGTTTGGCGATGAGAAAGACCTGCTGTAGAA 478

```

```

OY 116 SerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAla 135
Db 479 ACACATCGCTGGATGACAGCTTAAAGATATTTGGTACCTGACACCTTTTGTGA 538
OY 136 AsnGLysGlyAlaAsnPheHisGLuValThrThrPaspAsnLysLeuLeuArgIlePhe 155
Db 539 AATGAAAAACAATATTTTTCACATTGCAACAACCACTAATGATTCATCGTGTGAT 598
OY 156 LysAsnGLysAsnValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAsp 175
Db 599 CATCTGTGATCGATACAGAAAGATATTAGATTGACTAATACCGCATCTGTCGCGAT 658
OY 176 LeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGluSerSerIle 195
Db 659 CTACAAATTTTCCCATGATGCGCAGCTGTGCCACATTTGAATCGAA----- 706
OY 196 LeuYssSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuVal 215
Db 707 -----AGCTTGGTTACAGCATTCGAGATATCCGA 736
OY 216 PheGluTrpLeuGluAspAlaProAlaValGlnValGlnGlyLeuThrLeuProGln 235
Db 737 TATTTCTGGAGAGATGAGCTGAGTAGTGTGGCATGAGCAGTGAAGTGAACCTGAG 796
OY 236 PheIleLeuArgAspGluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLys 255
Db 797 TTCGAGTTTGGACACAGGACAG---AGGCGACCGCAAAATTAACCTAACCCAGGCAAC 853
OY 256 PheThrCysIleGluValLysPheHisLeuGlnArgGlnMetGlyTyrThrLeuIleGln 275
Db 854 TATTCGCGTTTACGCTGCGAATTCAGTCTGCTGGTGGATGAGGCTACTACTTATACAA 913
OY 276 MetTyrIleProSerLeuLeuValIleLeuSerTrpValSerPheTrpIleAsnMet 295
Db 914 ATCTACATACCCCTGCGACATGATCTGTATATCATGCGTATCATTTTGGCTCATGCG 973
OY 296 AspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrGln 315
Db 974 AATCAACGCGCGCGCGTGGCTGCTGCTGACCAACCGCTGTGACAAATGACCACTTTC 1033
OY 316 SerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrp 335
Db 1034 ATGCTGTCACAAATGCGCGCTGCCAAAGATTTCTGATCGTCAATTCATTTGAGCTCAT 1093
OY 336 MetAlaValCysLeuLeuPheValPheAlaAlaLeuGluGlyTyrAlaAlaIleAsnPhe 355
Db 1094 CTGGAGACATCTCTGTATGCTCTTCCAGTCACTAGCAATACGCCAGCGTGGCTAC 1153
OY 356 ValSerArgGlnHisLysGluPheIleArgLeuArgArgGlnArgGlnArgLeu 375
Db 1154 ATGGCAAAA-----CGAATTCAAATGCGAAAAACAAGATTTATGCGC 1195
OY 376 GluLysAspIleIleGlnGluSerArgPheThrPheArgGlyTyrGlyLeuGlyHisCys 395
Db 1196 ATCCAAAGATAGCGCAACAAAGACACAGCTCCAGGAGGAGGACCAACAGCAGCGC 1255
OY 396 -----LeuGlnAlaArgAspGlyLysProMetGluGlySerGlyIleTyrSerPro 412
Db 1256 AATCCCAATCCCAATGCAATATGTGGCGAGCC---GGAGAGATGGGGGTGGACCCGCG 1312
OY 413 GlnProProAlaPro 417
Db 1313 GGACCCGGAGGAGCC 1327

```

## RESULT 13

```

US-08-072-064-3
: Sequence 3, Application US/08072064
: Patent No. 6008046
: GENERAL INFORMATION:
: APPLICANT: FFERENCH-CONSTANT, RICHARD H.
: APPLICANT: JACKSON, MEYER B.
: TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING
: NUMBER OF SEQUENCES: 20

```

```

: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PETER G. CARROLL
: STREET: 220 Montgomery Street, Suite 2200
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC-compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/072.064
: FILING DATE: 19930602
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 770.881.
: FILING DATE: 04-OCT-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: CARROLL, PETER G.
: REGISTRATION NUMBER: 32.837
: REFERENCE/DOCKET NUMBER: OPHD-00574
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/705-8410
: TELEFAX: 415/397-8338
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2066 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: single
: TOPOLOGY: unknown
: MOLECULE TYPE: cDNA
: US-08-072-064-3

```

## Alignment Scores:

```

Pred. No.: 3.16e-82 Length: 2066
Score: 720.00 Matches: 155
Percent Similarity: 56.71% Conservative: 86
Best Local Similarity: 36.47% Mismatches: 144
Query Match: 31.97% Indels: 40
DB: 3 Gaps: 9

```

## US-10-075-846-4 (1-431) x US-08-072-064-3 (1-2066)

```

OY 5 ValProAlaThrLeuSerPheLeuLeuThrPheLeuProGlyGlnValLeuLeuArg 24
Db 137 CTGCCCGGCACAC---CCGCTCTTACCACTTGG-----CTGCCCATCAC 178
OY 25 ValAlaLeuAlaLysGluGlu-----ValLysSerGlyThr 36
Db 179 ATGGCCCTGATTGACAGAGAAAGCGGCCACAAACGATTCATACAGTACAGCAAGCGCGACT 238
OY 37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSer 56
Db 239 GCGCGTGGCACATCGTGGGTGACGTAAACCTATTCGCTTCTCGACTCCCTTATGCTT 298
OY 57 GlyTyrAspAlaArgIleArgPheProAsnPheLysGlyProProValAsnValThrCysAsn 76
Db 299 AGTTACGCAAAAGAGTAAGACCAATTACCGTGTCCCTGTGGAGGTTGGCGTCA 358
OY 77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnVal 96
Db 359 ATGTATGTCCTCACTATACATTCGTTTCCGAAGTCTCAATGCACTTCACATTTGATTTT 418
OY 97 PheLeuArgGlnGlnTrpPaspAspProArgLeuSerIleTyrArgGlyTyrProAsp---Asp 115
Db 419 TACTTTGCTCAATTTTGGACCGATCTCGTTTACCGTATAGAAACGACCTGTGTAGAA 478
OY 116 SerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPhePheAla 135
Db 479 ACACATCGCTGGATGACAGCTTAAAGATATTTGGTACCTGACACCTTTTGTGA 538

```

136 AsngluysgLyAlaAsnPhenHisGluValThrThrAspAsnLysLeuLeuArgIlePhe 155  
1539 AATGAAAAAACAATCATATTTCACATTCGACACACAGATTAAGATTCATACGTGCAT 598  
156 LysAsngLyAsnValLeuTyrSerIleArgIleLeuThrLeuLeuSerCysLeuMetLys 175  
599 CATTCGTGATCATTAACAAGAAAGATTAAGTACTATTAACCGCATCGCTCCGATGAT 658  
176 LeuLysAsnPhenProMetAspIleGlnThrCysThrMetGlnLeuGlySerSerIle 195  
659 CTACATATTTCCTCCATGATGATGCGACGCTGCCACATTTGAATGAA----- 706  
196 LeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAsnVal 215  
707 -----ACCTTCGCTTACCATGCGAATATCCGA 736  
216 PheGluThrLeuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrProGln 235  
737 TATTCCTGAGAGATGAGATGAGTGTGTCATGACAGACAGATGACATACCGCAG 796  
236 PheIleLeuArgAspGluLysAsnLeuGlyCysCysThrLysIleTyrAsnThrGlyLys 255  
797 TTCCGAGTTTGGGACACAGGCGAG---AGGCGCACCAGAAATTAACCTAACACAGCAAC 853  
256 PheThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGln 275  
854 TATTCGCGTTTACGCTGGGAATTCAGTTTCGTCGCTGATGCGCTACGCTTATACAA 913  
276 MetTyrIleProSerLeuLeuIleValIleLeuSerThrProValSerPheThrIleAsnMet 295  
914 ATCTCAATACCTCTGAGATGATGCTGTATATATCATGAGGTATCATTTTGGCTCAATCGC 973  
296 AspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThrGln 315  
974 AATGACAGCCGCGCGCTGTGCTGCTGCTGACACACGCTTGAACATGACCACTTGG 1033  
316 SerSerGlySerArgAlaSerLeuProLysValSerTyrValAlaIleAspIleThr 335  
1034 ATGTCGTCAACAATGACCGCTGTCGCAAAAGATTTGTCGTAACATGATGACGTAT 1093  
336 MetAlaValCysLeuLeuPheValPheAlaIleLeuGluTyrAlaIleAsnPhe 355  
1094 CTGGGACATGCTGCTGATATGCTCTTGTGCTACTGATGATGACCGCCGCTCGCTAC 1153  
356 ValSerArgGlnHisLysGluPheIleArgAlaArgArgGlnArgGlnArgLeu 375  
1154 ATGGCAAAA-----CGAATTCAAATCGCAAAACAAGATTTATGCGC 1195  
376 GluGluAspIleIleGlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlnHisCys 395  
1196 ATCCAAAAGATGCGGACAGAAAAGCAACAGCTCGACGAGCGAACCAACAGCAGCGC 1255  
396 -----LeuGlnAlaArgAspGlyIleProMetGlyIleSerGlyIleTyrSerPro 412  
1256 AATCCCAATCCCAATGCAATGCGGCGGAGCC---GAGAGAGTGGCGCTTGACCGCGC 1312  
413 GlnProProAlaPro 417  
1313 GGACCCGGAGGAGCC 1327

RESULT 14  
US-08-072-064-5  
Sequence 5, Application US/08072064  
Patent No. 6008046  
GENERAL INFORMATION:  
APPLICANT: FRENCH-CONSTANT, RICHARD H.  
APPLICANT: JACKSON, MEYER B.  
TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PETER G. CARROLL  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco

STATE: California  
COUNTRY: United States of America  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/072,064  
FILING DATE: 19930602  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 770,881  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: CARROLL, PETER G.  
REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: OPND-00574  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/705-8410  
TELEFAX: 415/397-8338  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2066 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-08-072-064-5

Alignment Scores:  
Pred. No.: 5 73e-82 Length: 2066  
Score: 718.00 Matches: 146  
Percent Similarity: 60.21% Conservative: 87  
Best Local Similarity: 37.73% Mismatches: 118  
Query Match: 31.88% Indels: 36  
DB: 3 Gaps: 8

US-10-075-846-4 (1-431) x US-08-072-064-5 (1-2066)

5 ValProAlaThrLeuSerPheLeuLeuLeuTyrThrLeuProGlyGlnValLeuLeuArg 24  
137 CTGCCCCGACACA---CCGCTGCTAACCATCTCG-----CTGGCATTCAAC 178  
25 ValAlaLeuAlaLysGlu-----ValLysSerGlyThr 36  
179 ATGGCCCTGATTGCACAGCAAGAGGCGCACAAACGATCATACAGTGCACACGCGGACT 238  
37 LysGlySerGlnProMetSerProSerAspPheLeuAspLysLeuMetCylArgThrSer 56  
239 GCGGCGGACGATGCTGGGTGACGTAACATATCCGCTATTCCTGACTCTTTAGTGT 298  
57 GlyTyrAspAlaArgIleArgProAsnPhenLysGlyProProValAsnValThrCysAsn 76  
299 AGTTACGACAAAAGATGAAGCAACCAATTACGCGTGTGCCCCCTGCGAGGTTGGCGTCACA 358  
77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnVal 96  
359 ATGTATGTCCTCAGTATCATGTCGTTTCGGAAGTTCTAATGACTTCACATTCGATTTT 418  
97 PheLeuArgGlnGlnIleTyrAsnAspProAlaGlySerTyrArgGlyTyrProAsp---Asp 115  
419 TACTTTCGTCATTTTGGACCGATCTCTTACGCTATGAAACAGACCTGCTGTAGAA 478  
116 SerLeuAspLeuAspProSerMetLeuAspSerIleTyrLysProAspLeuPheAla 135  
479 ACAGTATCGCTTGATGATCAAGTTCAATTAGATTTTGGGTACCGACACCTTTTGTGA 538  
136 AsngluysgLyAlaAsnPhenHisGluValThrThrAspAsnLysLeuLeuArgIlePhe 155  
539 AATGAAAAAACAATCATATTTCACATTCGACACACAGATTAAGATTCATACGTGCAT 598

[illegible]

```

      FILING DATE: 15-OCT-1993
      CLASSIFICATION: 435
      ATTORNEY/AGENT INFORMATION:
      NAME: Timian, Susan J.
      REGISTRATION NUMBER: 34,103
      REFERENCE/DOCKET NUMBER: 19603/120
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (716)263-1536
      TELEFAX: (716)263-1600
      INFORMATION FOR SEQ ID NO: 3:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1491 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: cDNA
      US-08-137-614A-3

Alignment Scores:
Pred. No.: 4.34e-82 Length: 1491
Percent: 717.00 Matches: 161
Score Similarity: 57.38% Conservative: 76
Best Local Similarity: 38.98% Mismatches: 125
Query Match: 31.84% Indels: 52
DB: 1 Gaps: 9

US-10-075-846-4 (1-431) x US-08-137-614A-3 (1-1491)
QY 1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuTyrPThrLeuProGlyGln 20
   |||||
Db 1 ATGCATCGTTTACGGCGCGTCGGAGATGCTTACCGCTGTTCTTTCTACTAGCGGCC 60
QY 21 ValLeuLeuArgValAlaLeuAlaValysGluGluValLysSerGlyThrLysGlySerGln 40
   |||
   ::::
   ::::
Db 61 CAGCTACAATGATGTTCGATGCATTCGAAAGAGATGACTACGCGCCGCTGAGAACGCTG 120
QY 41 ProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAla 60
   |||
   ::::
   ::::
Db 121 ACGCAACATATATCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 159
QY 61 ArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsn 80
   |||||
   ::::
   ::::
Db 160 CGACTTAGCGCCCAATTCGGAGAGACGACATGTCGGCGCATGATGATTCACCATGCC 219
QY 81 SerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGln 100
   |||||
   ::::
   ::::
Db 220 AGCTTGATGCCATATCAGAACAGTAAATCATGATATACGATAAACAATGATATTAATCAG 279
QY 101 GlnTyrAsnAspProArgLeuSerTyrArg-----GluTyrProAspAspSer--- 116
   |||
   |||
   |||
Db 280 TATTGGCCGACGACGAGGCTTTGGCATTTAATATCTTGGACAATATTCGACGATGAGAA 339
QY 117 -----LeuAspLeuAspProSerMetLeuAspSerIleTyrLys 129
   |||
   |||
   |||
Db 340 GATGATGGCATATAGCGATGTGCTGACATTTATCCGAGACTTTCGCGAAAGATATGGGTA 339
QY 130 ProAspLeuPhePheAlaAsnGluLysGlyValAsnPheHisGluValThrThrAspAsn 149
   |||||
   |||||
   |||||
Db 400 CCGGATACGTTCTTCCGCAATGACAAACAGATTTTTCGACAGATGTCACGGAAGAAC 459
QY 150 LysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIleArgLeuThrLeuIle 165
   |||||
   |||||
   |||||
Db 460 AAATGTGTGGACGTTGCGCGCGCATGGAGCTGTACTTTATGGCATAGAGATTCACGACACC 519
QY 170 LeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnInThrCysThrMetGln 189
   |||||
   |||||
   |||||
Db 520 CTGCGCTGCATGATGATCTGCATCTACTACTTCATTTGCACTGGCAGAAATGGACGTGGAA 579
QY 190 LeuGluSerSerIleLeuLysSerProLeuProSerLeuSerLeuSerValGlyTyr 209
   ::::
   ::::
   ::::
Db 580 ATTGAG-----ACGTATGAGTAC 597
QY 210 ThrMetLysAspLeuValPheGluTyrPheGluAsnAlaProAlaValGlnValAlaGlu 229

```

```
Db 598 ACGGTACGCGATGTGTCATGCTACTG---AAGCCAAAGCCAA---GTGCGCGAGATGAG 651
OY 230 GlyLeuThrLeuProGlnPheIleu-----ArgaspGluLysAspLeu 244
Db 652 GATCGGAGCGCGCGATTCACCATCATTTGGTATGAGACCAATGACCGAAAGAG--- 708
OY 245 GlyCysCysThrLysHisIleAsnThrGlyLysPheThrCysIleGluValLysPheHis 264
Db 709 -----CGGCTGGCCACTGAGCTATACGCGCTCTGCTCATTTCAA 753
OY 265 LeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuIleVal 284
Db 754 CTGCAACGGAATATCGGATCTTGTATTCCAACTATATCGCCAGCATTTCTGATCGTA 813
OY 285 IleLeuSerTrpValSerPheTrpIleAsnMetAspAlaIleProAlaArgValGlyLeu 304
Db 814 ATGCTGTGCTGGCTCTGCTCTGATTAACACGAGGCGAGAGTGCCTGCTCATTTG 873
OY 305 GlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuPro 324
Db 874 GGCATCACACGAGTGCCTCACCATGACCACTTAGCAGGGGTTCGACGCTCACTGCCG 933
OY 325 LysValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPhe 344
Db 934 CGCATATCGTATGGAAGGCGATCGACATTTATCTGTCATGTGCTCTGCTTTCGTGTC 993
OY 345 AlaAlaLeuGluGlyTyrAlaAlaIleAsnPhe-----ValSerArg 358
Db 994 GCACGCCCTTGGAATAGCGTCCGCTTAACATTACTTAAGGGCCAAAGGGCTAA-AA 1052
OY 359 GlnHisLysGlnPheIleArgLeuArgArgGlnArgGlnArgGlnArgLeuGluLasp 378
Db 1053 GAAATATAAGAAAGTCAAGAAATGTGTCCAGGCAAGATCGGAAAGAGTGAAGATCCGA 1112
OY 379 IleIleGlnGluSerArgPheTrpPheArgGlyTyrGly 391
Db 1113 GACGTGTTCACAGACAGACAGATTAATGACGCTGCAGGA 1151
```

Search completed: July 1, 2003, 00:57:28  
Job time : 115.6 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:07:05 ; Search time 229.481 Seconds  
(without alignments)  
2788.018 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252  
Sequence: 1 MTLVPATISFLILMTLPQ.....POPPAPILREGCTTRKLYVD 431

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=framet.p2n.model -DEV=x1p  
-O=/cgn2\_1/USPTO.spool/US10075846/runat\_25062003\_163649\_5181/app.query.fasta.1.782  
-DB=Published.Applications\_NA\_OPMT-fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsun62  
-TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10075846.ecgn.1.1.102\_etunat\_25062003\_163649\_5181  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGESOURCE -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published.Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2252	100.0	9	US-10-075-846-3
2	2163	96.0	9	US-10-075-846-1
3	1588	70.5	9	US-10-075-846-9
4	750	33.3	9	US-09-969-844-13

5	749	33.3	1609	9	US-09-969-844-11	Sequence 11, Appl
6	742.5	33.0	1197	9	US-10-239-420-10	Sequence 10, Appl
7	728.5	32.3	1640	9	US-09-969-844-12	Sequence 12, Appl
8	726.5	32.3	1467	10	US-09-815-925-10	Sequence 10, Appl
9	719	31.9	3442	9	US-10-239-420-3	Sequence 3, Appl1
10	719	31.9	3598	9	US-10-239-420-1	Sequence 1, Appl1
11	698.5	31.0	1693	9	US-10-037-270-131	Sequence 131, App
12	694.5	30.8	1866	10	US-09-964-824-96	Sequence 96, Appl
13	680.5	30.2	1555	9	US-10-211-673-11	Sequence 11, Appl
14	670	29.8	1150	9	US-10-239-420-13	Sequence 13, Appl
15	655	29.1	2194	9	US-10-239-420-4	Sequence 4, Appl1
16	652.5	29.0	1323	9	US-09-510-662A-38	Sequence 38, Appl
17	652.5	29.0	1323	10	US-09-778-320-38	Sequence 38, Appl
18	652.5	29.0	1323	12	US-10-010-742-38	Sequence 38, Appl
19	652.5	29.0	2593	9	US-10-106-698-5031	Sequence 2031, Ap
20	652.5	29.0	3282	9	US-09-954-531-1034	Sequence 1034, Ap
21	652.5	29.0	3282	10	US-09-825-301-75	Sequence 75, Appl
22	652.5	29.0	3282	10	US-10-239-420-6	Sequence 6, Appl1
23	640.5	28.4	4077	10	US-09-808-483-5	Sequence 5, Appl1
24	632.5	28.1	1192	10	US-09-808-483-3	Sequence 3, Appl1
25	632.5	28.1	1269	10	US-09-808-483-7	Sequence 7, Appl1
26	632.5	28.1	1272	10	US-10-037-270-256	Sequence 256, App
27	625	27.8	1997	9	US-09-818-657-1	Sequence 1, Appl1
28	620	27.5	2819	9	US-10-211-673-7	Sequence 7, Appl1
29	619.5	27.5	1707	9	US-10-037-270-331	Sequence 331, App
30	619	27.4	1987	10	US-09-893-321-5	Sequence 5, Appl1
31	616	27.4	1398	10	US-09-893-321-1	Sequence 1, Appl1
32	616	27.4	1568	10	US-09-808-483-11	Sequence 11, Appl
33	615	27.3	1458	10	US-09-808-483-9	Sequence 9, Appl1
34	615	27.3	1608	10	US-09-808-483-1	Sequence 1, Appl1
35	608	27.0	1404	9	US-09-765-069-1	Sequence 1, Appl1
36	608	27.0	1422	10	US-09-742-311-1	Sequence 1, Appl1
37	607	27.0	1263	9	US-09-765-069-7	Sequence 7, Appl1
38	603	26.8	1179	9	US-09-765-069-3	Sequence 3, Appl1
39	602	26.7	1038	9	US-09-765-069-9	Sequence 9, Appl1
40	594.5	26.4	1444	9	US-09-839-446-5	Sequence 5, Appl1
41	594.5	26.4	1444	10	US-09-898-570-5	Sequence 5, Appl1
42	592.5	26.3	1438	9	US-09-839-446-7	Sequence 7, Appl1
43	592.5	26.3	1438	10	US-09-898-570-7	Sequence 7, Appl1
44	591	26.2	1417	9	US-09-839-446-3	Sequence 3, Appl1
45	591	26.2	1417	10	US-09-898-570-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-075-846-3  
Sequence 3, Application US/10075846  
Publication No. US20030032608A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRN4, and SPLICE VARIANT THERE  
FILE REFERENCE: D0079 NP  
CURRENT APPLICATION NUMBER: US/10/075, 846  
CURRENT FILING DATE: 2002-02-13  
PRIOR APPLICATION NUMBER: US 60/269, 535  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 1640  
TYPE: DNA  
ORGANISM: homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1293)  
US-10-075-846-3  
Alignment Scores:  
Pred. No.: 1.62e-278  
Score: 2252.00  
Length: 1640  
Matches: 431  
Percent Similarity: 100.00%  
Conservative: 0





Db 241 AGTTTCAGCTCCGTCACCAAGACCACCAATGACCTACCGGGTGAATGCTCTTGGCGCAA 300  
Qy 101 GlnTPrAsnAspProArgLeuSerTyrArgGluTyrProAspSpsSerLeuAspLeuAsp 120  
Db 301 CAGTGAAATACACCGCCCTGCTTACCGAGAAATATCTGATACCTCTGCGACCTCGAT 360  
Qy 121 ProSerMetLeuAspSerIleTPrLysProAspLeuPhePheIleAsnGluLysGluVal 140  
Db 361 CCTCCATGCTGACCTATCTGGAAGCCAGACCTCTCTTCTGCTAATGAGAAAGGGGCC 420  
Qy 141 AsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGluVal 160  
Db 421 AACTTCATGAGGTGACACGAGCAACAACTTACTGGCATCTTCAAGATGGGAATGTG 480  
Qy 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetSpleuLysAsnPhePro 180  
Db 481 CTGTACAGCTACAGGCTGACCTCATTTTGTCTGCTGTATGGACCTCAAGAACTTCCCC 540  
Qy 181 MetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeu 200  
Db 541 ATGACATCCAGACCTGACGATGCAAGCTTGAG-----573  
Qy 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTPrLeuGlu 220  
Db 574 -----AGCTTTGGCTACACCATGAAGACCTCGTGTGGAGTGGCTGGAA 618  
Qy 221 AspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAsp 240  
Db 619 GATGCTCTCTGCTGCACAGTGGCTGAGGGGCTGACTGCTGCCAGTTATCTTGGCGGAT 678  
Qy 241 GlnLysAspLeuGlyCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu 260  
Db 679 GAGAGAGATCTAGGCTGTGTACCAACGACTACACAGCAGGAATTCACCTGCATGAG 738  
Qy 261 ValLysPheHisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSer 280  
Db 739 GTAAGATTTACCTGGAACGCGAGATGGGCTACTATGATTCAGATGATCATCCCGAC 798  
Qy 281 LeuLeuIleValIleLeuSerTPrValSerPheTPrIleAsnMetAspAlaIleProAla 300  
Db 799 CTACTCATCTCATCCCTGCTGCTGGGTCTCTTCTGATCAACATGATGCTGCCCTGCC 858  
Qy 301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArg 320  
Db 859 CGGTGGGCTGGGACATCACACCGTCTCACCATGACCAACGAGCTGGGCTCCCGG 918  
Qy 321 AlSerLeuProLysValSerTyrValLysAlaIleAspIleTPrMetAlaValCysLeu 340  
Db 919 GCCTCTTTGGCTTAAGGTGCTCTACGTGAAGCAATCGCATCTGGATGCTGTGCTCTG 978  
Qy 341 LeuPheValPheAlaIleLeuGluTyrValAlaIleAsnPheValSerArgGlnHis 360  
Db 979 CTCTTTGTCTGCTGCTGCTGGGTGAGTATGCTGCCATTAATTTTGTCTTCTGCTGAC 1038  
Qy 361 LysGlnPheIleArgLeuArgArgGlnArgGlnArgGlnArgGlnArgGlnArgGln 380  
Db 1039 AAGAAATTCATACGACTTCGAAGAAGCAGAGGCGCAACGCTTGAGGAATATCATC 1098  
Qy 381 GlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlnLysCysLeuGlnAlaArgAsp 400  
Db 1099 CAAGAAAGTGGTTTATTTTCCGGCTATGGCTTGGGCCATCTCCGACGAGGAAGAT 1158  
Qy 401 GlyGlyProMetGluGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg 420  
Db 1159 GGAAGTCCATGGAAGTCTTGGCATTTATAGTCCCAACCTCCAGCCCTTCTTAAGG 1218  
Qy 421 GlnGlyGluThrThrArgLysLeuTyrValAsp 431  
Db 1219 GAAGGAGAAACACGCGGAATCTACGTGAC 1251

RESULT 3  
US-10-075-846-9  
; Sequence 9, Application US/10075846

; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HEPAR4, and SPLICE VARIANT THER  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 993  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; US-10-075-846-9  
  
Alignment Scores:  
Pred. No.: 1,29e-193 Length: 993  
Score: 1588.00 Matches: 309  
Percent Similarity: 89.86% Conservative: 1  
Best Local Similarity: 89.57% Mismatches: 1  
Query Match: 70.52% Indels: 34  
Gaps: 2  
DB:  
  
US-10-075-846-4 (1-431) x US-10-075-846-9 (1-993)  
Qy 1 MetThrThrLeuValProAlaThrLeuSerPheLeuLeuLeuThrThrLeuProGlyGln 20  
Db 1 ATGACAACTCTGTTCCTCGCAACCTCTCTCTTCTTCTGAGACCCGCGAGGGCAG 60  
Qy 21 ValLeuLeuArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGln 40  
Db 61 GTCCCTCTCAGGGGCTTGGCGCAAAAGAGAGTCAAAATCTGAACCAAGAGGGTCCAG 120  
Qy 41 ProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAla 60  
Db 121 CCCATGCCCCCTCTGATTTCTTAGACAACTTATGGGGCAACATCTGATATGATGCC 180  
Qy 61 ArgIleArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsn 80  
Db 181 AGGATTCGGCCCAATTTTAAAGGCCCCACCGTGACGTGACATGCAACATCTTCATCAC 240  
Qy 81 SerPheSerValThrLysThrThrMet-----90  
Db 241 AGTTTCAGCTCCATCCCAAGACCACAAATGGCTTGGGCCCCCTGGGAATGCCATGTT 300  
Qy 91 -----AspTyrArgValAsnValPheLeuArgGln 100  
Db 301 TCTGAAGGGCCCATATCTGCACCCCTCCAGAGCTACCGGGTGAATGCTTCTTGGCGCAA 360  
Qy 101 GlnTPrAsnAspProArgLeuSerTyrArgGluTyrProAspSpsSerLeuAspLeuAsp 120  
Db 361 CAGTGAAATACACCGCCCTGCTTACCGAGAAATATCTGATGACTCTTCTGGAACCTCGAT 420  
Qy 121 ProSerMetLeuAspSerIleTPrLysProAspLeuPhePheIleAsnGluLysGluVal 140  
Db 421 CCTCCATGCTGACCTATCTGGAAGCCAGACCTCTTCTTCTGCTAATGAGAAAGGGGCC 480  
Qy 141 AsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGluVal 160  
Db 481 AACTTCATGAGGTGACACGAGCAACAACTTACTGGCATCTTCAAGATGGGAATGTG 540  
Qy 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetSpleuLysAsnPhePro 180  
Db 541 CTGTACAGCTACAGGCTGACCTCATTTTGTCTGCTGTATGGACCTCAAGAACTTCCCC 600  
Qy 181 MetAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerProLeu 200  
Db 601 ATGACATCCAGACCTGACGATGCAAGCTTGAG-----633  
Qy 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTPrLeuGlu 220





Db	12	AGATTGATGACCTCCACCATGAGTACACAGTGCACAAATGACGTTCCAGAGAGCAGTGGCGG	180
Oy	104	AspProArgLeuSerTyrArgLutYrProAspSerLeuAspLeuAspProSerMet	123
Db	181	GAGCAGAGACTCCAGTACGACGACTTGGGGCGCCAGTTGCTACTACGACCTTACCAGAA	240
Oy	124	LeuAspSerIleTyrPryProAspLeuPhePheIleAsnGluTyrSGlyAlaAsnPheHis	143
Db	241	CGGACAGACGTTTGGAAAGCGGACCTGTTTTCTCTCCAGAGAAAGAGGGACACTTCCAC	300
Oy	144	GluValIleThrAspAsnLysLeuLeuArgIlePheIlySaSnGlyAsnValLeuTyrSer	163
Db	301	ACACATCATGATGCCCAACCTGCTTACGACATACATCCCAACGGCGACGCTTCTTCAGC	360
Oy	164	IleArgLeuThrIleuLeuSerCysLeuMetLaspLeuLaspLeuLaspLeuPhePrometAspIle	183
Db	361	ATAGAGATATCTTGTTGGCTTTCATGTCGAGTGAACCTGGAATTTTATCTTGTGATPAA	420
Oy	184	GlnThrCysThrMetGlnLeuGlnIleuSerSerIleLeuCysSerProLeuProSerLeu	203
Db	421	CAATCTGCTCTATC-----GTC	438
Oy	204	SerLeuSerValGlyIlyrThrMetLysAspLeuValPheGluTyrPleuGluAspPalaPro	223
Db	439	ATGGTAGCGATGCGATGATCAACAGACGCCGCGTTCATCATGAAAGAGGGGAGTCTT	498
Oy	224	AlaValGlnValIaGluLeuThrLeuThrLeuProGlnPheIleLeu---ArgAspGluLys	242
Db	499	---GTACAGGTCCAAAAATCTCCACACTTCCACGCTTACAGCTCGGAAGAGTTTCAAC	555
Oy	243	AspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLys	262
Db	556	GACTAC-----TGCCACCGATGGGACCAACATGGCGAGTACACTGCTGGCGGTGAC	609
Oy	263	PheHisLeuGlnArgGlnMetGlyTyrTyrIleuIleGlnMetTyrIleProSerLeuLeu	282
Db	610	CTGGTGTTCACAGCGCAGATTACAGTACATACCTGATCCAGATCTACATCCGCTGTCATG	669
Oy	283	IleValIleLeuSerTyrValSerPheThrPileasMetLaspAlaIaProIaIaArgVal	302
Db	670	CTGGTCATGCTGTCCTGGGCTGTCGTTTGGCTGCACCCACCTGATGCCGCGCGAGTG	729
Oy	303	GlyLeuGlyIleThrThrValLeuThrMetThrThrGlnIleSerArgIaIaSer	322
Db	730	TCGCTGGGCGCTCACCCCTGCTCACCATGGCCACGAGATATGGGAGATCACGCTCG	789
Oy	323	LeuProLysValSerTyrValLysAlaIleAspIleTyrPheIaValaCysLeuLeuPhe	342
Db	790	CTGCTCCGCTTCTTCTACACCAAGGCGCATTCAGTGGACCGGCGTCTGTCTGCACTTC	849
Oy	343	ValPheIaIaIaLeuLeuGluTyrAlaIaIaIeAsnPheValSerATyrGlnHisLysGlu	362
Db	850	GTAATGGGGCGCTCTCTGAGTTGCCCTGGTGCACATCAGCGCTCGGATTCACGC	909
Oy	363	PheIleArgLeuArgArgArgGlnArgGln-----ArgLeuGlnGlu	377
Db	910	CGGACAGACATGCAAGACAGAGAGAGAGAAATGGAGCTCAGACCCGCCCTTGACTCG	969
Oy	378	AspIleIleGlnGluSerArgPheTyrPheArgGlyTyrGlyLeuGlyHis	394
Db	970	GACCACCTGGAGAGCAGCGCCACACAGTTCGCGCATGAGAGCGCGTGTGTCAC	1020
RESULT 7			
US-09-969-844-12			
; Sequence 12, Application US/09969844			
; Publication No. US20020192776A1			
; GENERAL INFORMATION:			
; APPLICANT: Xiao-Zhou Michelle Wang			
; APPLICANT: Xavier Georges Sarda			
; APPLICANT: Michael David Tomalski			
; APPLICANT: Vincent Paul Mary Wingate			
; TITLE OF INVENTION: Heliotris Glutamate Receptor			
; FILE REFERENCE: A32815-1 072657.0178			

```

CURRENT APPLICATION NUMBER: US/09/969,844
CURRENT FILING DATE: 2001-10-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 12
LENGTH: 1640
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: PCR amplified fragment of Helicobis DNA cloned
US-09-969-844-12

Alignment Scores:
Pred. No.: 7,06e-83 Length: 1640
Percent: 728.50 Matches: 156
Score: 61.108 Conservative: 56
Best Local Similarity: 44.96% Mismatches: 110
Query Match: 32.35% Indels: 25
DB: 9 Gaps: 8

US-10-075-846-4 (1-431) x US-09-969-844-12 (1-1640)

QY 31 GluValLysSerGlyThrLysGlySerGlnPrometSerProSerAspPheLeuAspLys 50
Db 154 GAATTCATGACGATCGGAGAAATCAACATCTTCAGAGAGAGAGAGAGACGATCTGGATCAG 213
QY 51 LeuMetGlyArgThrSerGlyThrAspAlaArgLeuArgPro-----AspPheLys 67
Db 214 ATCTGGGC---CCCGGAGGTACGACGCGCAGATCAGACCTCGGGATCAGCGCACT 270
QY 68 Gly--ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThr 86
Db 271 GGCATATCCGCAACGTTAGTCATGCAACATGATATCTACGTCATCAGCAAAATGAT 330
QY 87 LysThrThrMetAspIrrArgValAsnValPheLeuArgGlnGlnThrPAsnAspProArg 106
Db 331 GATTCAAAATGGAAATACCTCCGATTAACGTTTCGGGAACAATGGTTAGTAAGACGG 390
QY 107 LeuSerIrrArgLysIrrProAspAspSerLeuAspLeuAspProSerMetLeuAspSer 126
Db 391 CTCGAATTCATATATCTTGAGAGTGGCCCAATACCTGACATGACATGCAAGCCACAGA 450
QY 127 IleThrLysProAspLeuPhePheAlaAsnGlnLysGlyAlaAsnPheHisGluValThr 146
Db 451 GTCGGATGCGCTATCTATCTTCCCAACGAGGAAGAGTCTTCCACAAACATCATC 510
QY 147 ThrAspAsnLeuLeuArgIlePheLysAsnGlnAsnValLeuIrrSerIleArgLeu 166
Db 511 ATGCCAAGCTGTACATCCGATCTTCCCAACGCAACGTCGCTGTACAGATCCGATC 570
QY 167 ThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePrometAspIleGlnThrCys 186
Db 571 TCCCTGACGCTCTCGTCCGCCCATGAACCTCAAGTTGATCCCTCGATTAAGCAACCTGC 630
QY 187 ThrMetGlnLeuLeuSerSerIleLeuCysSerProLeuProSerLeuSerLeuSer 206
Db 631 TCCGTCAGCATG-----GCTAGT 648
QY 207 ValGlyIrrThrMetLysAspLeuValPheGluIrrPheGlnGluAspAlaProAlaValGln 226
Db 649 TATGTTGGACCAACAGACGATTAAGTTCCTATGGAAGGAAGGCAACCCG--GTCCAG 705
QY 227 ValAlaGlnLysLeuThrLeuProGlnPheIleLeuArgAspGlnLysAspLeu--Gly 245
Db 706 GTGTGAAAAAATTTACACCTCGCTGGTTTCACGCG-----GAGAAGTTCTTACTAGAC 759
QY 246 CysCysThrLysHisIrrAsnThrGlnLysPheThrCysIleGluValLysPheHisLeu 265
Db 760 TACTGCAACAGTAAGACTAATACCGGTGAATACATTCCTCGAAGGTAGACCTGCTTC 819
QY 266 GluArgGlnMetGlyIrrThrLeuIleGlnMetIrrIleProSerLeuIleValIle 285

```

Db 820 AACGGAGTTCAGTACTACCTGATCCAGATCTACATTCCTGCTGCATGCTGTCATC 879  
Oy 286 LeuSerTrpValSerPheTrpIleAsnMetAspAlaIleProIleArgValIleGly 305  
Db 880 GTGTCTGGGTGCTCTGCTGCTGACCCAGGAGCTGTGCTCGAGGGCTCTCAGTACGA 939  
Oy 306 IleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLys 325  
Db 940 GTACAGCCTTACTTACAAATGAGGGACCCATGCTCAGGCAATCAACGGCTCCCTACCA 999  
Oy 326 ValSerTrpValIleAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAla 345  
Db 1000 GTGTCTTACACGAAAGCCATTGATGCTGACTGCTGATGCTCATTCATTCGTA 1059  
Oy 346 AlaleuLeuGlnIleValIleAlaIleAsnPheValSerArg-----GlnHisLysGluPhe 363  
Db 1060 GCCCTCTAGAGTTGTGGCTGCTGCACTATGCTGCTGCTGATCATGACGAGAGAAC 1119  
Oy 364 IleArgLeuArgArgArgGln 370  
Db 1120 ATGAGAAACGAGACGAGAG 1140

## RESULT 8

US-09-815-925-10  
: Sequence 10, Application US/09815925  
: Patent No. US20020127199A1  
: GENERAL INFORMATION:  
: APPLICANT: Tang, Y. Tom  
: APPLICANT: Zhou, Ping  
: APPLICANT: Goodrich, Ryle  
: APPLICANT: Asundi, Vinod  
: APPLICANT: Yang, Yonghong  
: APPLICANT: Zhang, Jie  
: APPLICANT: Mehrman, Tom  
: APPLICANT: Dimauc, Radoje T.  
: TITLE OF INVENTION: No. US20020127199A1el Nucleic Acids and  
: FILE REFERENCE: 787CIP2H  
: CURRENT APPLICATION NUMBER: US/09/815, 925  
: PRIOR FILING DATE: 2001-03-22  
: PRIOR APPLICATION NUMBER: 09/560, 875  
: PRIOR FILING DATE: 2000-04-27  
: PRIOR APPLICATION NUMBER: 09/496, 914  
: PRIOR FILING DATE: 2000-02-03  
: NUMBER OF SEQ ID NOS: 11  
: SOFTWARE: PL-FL-Genes Version 2.0  
: SEQ ID NO 10  
: LENGTH: 1467  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (46)..(1464)  
: US-09-815-925-10

## Alignment Scores:

Pred. No.: 1,06e-82 Length: 1467  
Score: 726.50 Matches: 157  
Percent Similarity: 58.59% Conservative: 68  
Best Local Similarity: 40.89% Mismatches: 111  
Query Match: 32,26% Indels: 48  
DB: 10 Gaps: 7

US-10-075-846-4 (1-431) x US-09-815-925-10 (1-1467)

Oy 4 LeuValProIleAlaThrLeuSerPheLeuLeuTrpThrLeuProGlyGlnValLeuLeu 23  
Db 58 CTCCTGGAGCTCCTGCTGCCATCTGCTCCTGGAACCTG----- 99  
Oy 24 ArgValAlaLeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnProMetSer 43  
Db 100 -----GGACCCGAGAGCTCTGAGCCCGCAGT 126

Oy 44 ProSerAsp-----PheLeuAspLysLeuMetGlyArg 54  
Db 127 GTACAGCATCCCGGAACATGTCCTTTGTGAAGAGACGCTGGACAAGCTGTG----- 180  
Oy 55 ThrSerGlyTrpAspAlaArgIleArgProAsnPheLysGlyProProValAsnValThr 74  
Db 181 ---AAAGGCTACAGACATTCGCTTAAGACCCGACTTGGGGGTCGCCGCTCGTGGGG 237  
Oy 75 CysAsnIlePheIleAsnSerPheSerSerValThrLysThrThrMetLysPyrArgVal 94  
Db 238 ATGACATGCACATCCACCATCGACATGACATGCTTCCGAAAGTCAACATGATGATTAACCTTA 297  
Oy 95 AsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSerTrpArgLysProAsp 114  
Db 298 ACCATGATTTTCAACAAATTTTGGACAGATAAAGGCTGCCATTTCTGGGATCCCT---- 354  
Oy 115 AspSerLeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPhePhe 134  
Db 355 CTCACCTCACGCTTGACATCGAGTGGCTGACACACTATGAGTGGCCGACACATATTTTC 414  
Oy 135 AlaAsnGlnLysGlyAlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIle 154  
Db 415 TTAATGACAAAAGTCATTGTCATGAGTACAGTGAAGAAACCGCATGATGCTCTT 474  
Oy 155 PheLysAsnGlyAsnValLeuTrpSerIleArgLeuThrLeuIleLeuSerCysLeuMet 174  
Db 475 CACCCTGATGGACAGTGTGCTGATGGCTGAGATACACCGACGACACGATCATGATG 534  
Oy 175 AspLeuLysAsnPhePheMetLeuAspIleGlnThrCysThrMetGlnLeuGlnSerSer 194  
Db 535 GACCTCAGAGATACCCCTGAGACGAGAACATGCACTCTGGAATTTGAA----- 585  
Oy 195 IleLeuCysSerProLeuProSerLeuSerLeuSerValGlyTrpThrMetLysAspLeu 214  
Db 586 -----AGCTATGCTTACACACACGATGACATT 612  
Oy 215 ValPheGluTrpLeuGluAspAlaProAlaValGlnValAlaGluGlnLeuThrLeuPro 224  
Db 613 GAGTTTACTGGGAGGAGGAGGAGCAAGGCTGTACCGAGTGGAAAGATGAGTCCCG 672  
Oy 235 GlnPheIleLeuArgAspCysLysAspLeuGlyCysCysThrLysThrLysAsnThrGly 254  
Db 673 CACTTCTCATGCTGAGCACCT---CTGCTCGAGGAATGTGCTTCCGCCACAGGT 729  
Oy 255 LysPheThrCysIleGluValLysPheHisLeuGlnArgGlnMetGlyTrpLeuIle 274  
Db 730 GCCTATCTCTGACTGCTGACCTGCTTGGTTGAAGAAACATTTGATTCATTCATT 789  
Oy 275 GlnMetTrpIleProSerLeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsn 294  
Db 790 CAGACTTATATGCCCCTCTATACGATACGATTTCTGTGGGTGCTCTTGTGATCAAT 849  
Oy 295 MetAspAlaIleProIleAlaArgValIleGlyIleThrThrValLeuThrMetThrThr 314  
Db 850 TATGATGATCATCTGCTAGATGTCCTGCGGATCACAACCTGCTGCAATGACAAC 909  
Oy 315 GlnSerSerGlySerArgAlaSerLeuProLysValSerTrpValLysAlaIleAspIle 334  
Db 910 ATCAACACCCACCTTGGGAGACCTTGGCCCAAAATCCCATATGCAAAAGCATTTGACATG 969  
Oy 335 TrpMetAlaValCysLeuLeuPheValPheAlaIleLeuGlnIleValIleAsn 354  
Db 970 TACCTTATGGGCTGCTTGTGCTTGTGCTTGTGCTTGTGAGATGCTTGTGCAAC 1029  
Oy 355 PheVal-SerArgGlnHis-----LysGluPheIleArgLeuArgArgGlnArgArg 372  
Db 1030 TACATTTTCTTGTGAGAGAGCCCTCAAGACGAAGAAGATTCGACAAAAGACGCAAG 1089  
Oy 372 gGlnArgLeu 375  
Db 1090 GCAAGAAGT 1099

## RESULT 9

```
US-10-239-420-3
: Sequence 3, Application US/10239420
: Publication No. US20030096984A1
: GENERAL INFORMATION:
: APPLICANT: Cully, Doris F.
: APPLICANT: Zheng, Yiqiong
: TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
: FILE REFERENCE: 20629P
: CURRENT APPLICATION NUMBER: US/10/239,420
: PRIORITY FILING DATE: 2002-09-23
: PRIOR APPLICATION NUMBER: PCT/US01/09956
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: 60/193,935
: PRIOR FILING DATE: 2000-03-31
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 3442
: TYPE: DNA
: ORGANISM: Dermacentor variabilis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (32)...(1225)
: US-10-239-420-3

Alignment Scores:
Pred. No.: 4.08e-81 Length: 3442
Score: 719.00 Matches: 150
Percent Similarity: 62.01% Conservative: 72
Best Local Similarity: 41.90% Mismatches: 98
Query Match: 31.93% Indels: 38
DB: 9 Gaps: 10

US-10-075-846-4 (1-431) x US-10-239-420-3 (1-3442)

QY 9 LeuSerPheLeuLeuLeuTrp-----ThrLeuProGlyGlnValLeuLeuArg 24
DB 38 CTTTCAGCGCTGCAACGCTGCGCGCTTGCCTCAGCTTGCTC-----CTCCTCAGG 88
QY 25 ValAlaLeuAlaValysGluValysSerGlyThrValysGlnPrometSerPro 44
DB 89 ACGAGCGCTCGCGCAAGGAGCGGCTCAACGAGCG----- 124
QY 45 SerAspPheLeuAspLysLeuMetCylarGThrSerGlyTyraSpAlaArgIleArpPro 64
DB 125 CTGATGACGCTGAGAACCTTGACGACTTATTAGAACCTTACACCGCGCGCTTCC 184
QY 65 ---AspPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83
DB 185 ACGACACACTTGGAACCGCAACAACAAAGTGGCTTGCCTTACATCGCAGACTTCGGG 244
QY 84 SerValThrLysThrMetAspTyrArgValAsnValPheLeuArgGlnIleTrpAsn 103
DB 245 TCCATTAATCCAGCCACACATGAGCTATGAGTTGATCTTTATTTCGGCAGACTTCGCA 304
QY 104 AspProAlaLeuSerTyrArgGluTyrProAspAspSer-----LeuAspLeu---Asp 120
DB 305 GATGATCGCTTGAGC-----AGCCCAACGATATCAGGCCCTGAGCTCAATGAT 355
QY 121 ProSerMetLeuAspSerIleTyrLysProAspLeuPhePheIleAsnGlnLysGlyAla 140
DB 356 CCAAGAGCTGCGACCGCTATATGGAACAGATATCTTCGCAAAATGCAAAACACCGCA 415
QY 141 AsnPheHisGluValThrThrAspAsnLysLeuArgIlePheLysAsnGlyAsnVal 160
DB 416 GAGTTCCAAATATGTCACAGTACCTTAATGTTACTGGTCCGCTTAACCCAGGAAAGATT 475
QY 161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePro 180
DB 476 CTATTCATGCTCAGGCTCAAGCTTAAGCTTTCATGATGATGATGATGATGATGATGAT 535
QY 181 MetAspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuCysSerProLeu 200
```

```
DB 536 ATGACACTCCCAAGTTTGACGATCGAACTCCCTCATTTCTCG----- 577
QY 201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTrpLeuGlu 220
DB 578 -----AAAAACAACCAAGAACGCTGCTGAGTGGTCCGAT 613
QY 221 AspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAsp 240
DB 614 ACCAATCCGATTAATA---CTATTGCAAGGCTGAGAGTTACACAGTTCCGAGATTACAGAT 670
QY 241 GluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu 260
DB 671 ---ATAAATACGTCATCTGCATGAGAAATTTTCATCGCAGAGTACAGCTGCTGAG 727
QY 261 ValLysPheHisLeuGluArgGlnMetGlyTyrTyrIleLeuIleGlnMetCylProSer 280
DB 728 GCCGACTTCCACTTGCGACGGCTCAGCGCTACACATGGTGCTGATCTGCTTCCATCA 787
QY 281 LeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAla 300
DB 788 GTGCTCATCGTGTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 847
QY 301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArg 320
DB 848 CGCACACACTGCGCGCTCAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 907
QY 321 AlaSerLeuProLysValSerTyrValLysAlaIleAspIleTrpMetAlaValCysLeu 340
DB 908 TCCACCTTGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 967
QY 341 LeuPheValPheAlaAlaLeuLeuGluTyrAlaIleAlaIleAsnPheValSerArg 358
DB 968 GCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
```

```
RESULT 10
US-10-239-420-1
: Sequence 1, Application US/10239420
: Publication No. US20030096984A1
: GENERAL INFORMATION:
: APPLICANT: Cully, Doris F.
: APPLICANT: Zheng, Yiqiong
: TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
: FILE REFERENCE: 20629P
: CURRENT APPLICATION NUMBER: US/10/239,420
: PRIORITY FILING DATE: 2002-09-23
: PRIOR APPLICATION NUMBER: PCT/US01/09956
: PRIOR FILING DATE: 2001-03-28
: PRIOR APPLICATION NUMBER: 60/193,935
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 3598
: TYPE: DNA
: ORGANISM: Dermacentor variabilis
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (170)...(1363)
: US-10-239-420-1

Alignment Scores:
Pred. No.: 4.39e-81 Length: 3598
Score: 719.00 Matches: 150
Percent Similarity: 62.01% Conservative: 72
Best Local Similarity: 41.90% Mismatches: 98
Query Match: 31.93% Indels: 38
DB: 9 Gaps: 10

US-10-075-846-4 (1-431) x US-10-239-420-1 (1-3598)

QY 9 LeuSerPheLeuLeuLeuTrp-----ThrLeuProGlyGlnValLeuLeuArg 24
```

```

Db      176 CTTTGACCCCTGAACGTGGGCGGCTTGCCTTCACCTTGTCC-----CTCCACAG 226
Oy      25 ValAlaLeuAlaLysGluValLysSerGlyThrLysGlySerGlnPrometSerPro 44
Db      227 ACGACCGCTGGCGAGAAAGCGGTCMAACGGAGCG----- 262
Oy      45 SerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLeuArgPro 64
Db      263 CTGGATGACCTGAGAAAGCTTGAAGCACTTATTAAGCAACCGCGGCTCCCTCC 322
Oy      65 ---AspPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83
Db      323 ACGACACACTTGGGAACGCCAACAAAAGGCTTGGCAATCTACATACGACCTTCGGG 382
Oy      84 SerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTyrPasn 103
Db      383 TCCATTAATCCAGCCCAATGACATGACTATGAGGTTGATCTTATTGGCGGACACTTGGCAA 442
Oy      104 AspProArgLysSerTyrArgGluTyrProAspAspSer-----LeuAspLeu---Asp 120
Db      443 GATGATCGGCTGACG-----ACCCCAACGATTCACAGCGCCCTGGACCTCATGAT 493
Oy      121 ProSerMetLeuAspSerIleTribLysProAspLeuPhePheAlaAsnGluLysGlyAla 140
Db      494 CCAAACTGCGTGGCGGCTATGAGAAACCGGAAGTATCTCCCAATGCCAAACACGCA 553
Oy      141 AsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnVal 160
Db      554 GACTTCCAATATGCTCACTACCTAATGTATGCTGCTGCTTAAACCCGGAAGAAAGT 613
Oy      161 LeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePro 180
Db      614 CTATACATGCTCGTCAAGCTCAAGTTTGCATGTATGATGATGATTTATTCCTTCC 673
Oy      181 MetAspIleGlnThrCysThrMetGlnLeuGlnSerSerSerIleLeuLysSerProLeu 200
Db      674 ATGGACTCCCAAGTTCGACACGACGAACTGCGCTCATTCGCG----- 715
Oy      201 ProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrPleuGlu 220
Db      716 -----AAACACACCGAAAGAACTGCATCTGGAGTGTCTGAT 751
Oy      221 AspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAsp 240
Db      752 ACCAATCCGATATA---CTATTCGAAGCGCTGAGTTACCAATTCGAGATTTCAGAA 808
Oy      241 GluLysAspLeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGlu 260
Db      809 ---ATAAATACGTCAATCTGCATGAGAAATTCACATCGAGACAGTACACTCCCTGAG 865
Oy      261 ValLysPheHisLeuGlnArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSer 280
Db      866 GCGGACTTCACCTGACGCGGTGCTACGCGGTACCAATGTCGATGTCGATTCGCTACA 925
Oy      281 LeuLeuIleValIleLeuSerTyrValSerPheTyrIleAsnMetAspAlaAlaProAla 300
Db      926 GTGCTCATGCTGTCATCTGCGGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
Oy      301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrIleSerSerGlySerArg 320
Db      986 CGGACACACACTGGCGCTGACGAGCGCTCACAATTCCTCCAAAGGCTCCGCTATACAG 1045
Oy      321 AlaSerLeuProLysValSerTyrValLysAlaIleAspIleThrMetAlaValCysLeu 340
Db      1046 TCCAACTGCTGCTCGGCTCATACGTCGAAGCAATGATGTGTGATGGAGCTGCACG 1105
Oy      341 LeuPheValPheAlaLeuLeuGluTyrAlaAlaIleAsnPheValSerArg 358
Db      1106 GCGCTGTGTTCTTCGCGACACTAGGATTCACCGCTGCTGCTGCTGCTGCTGCTGCTG 1159

```

```

; Sequence 131, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundui
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dramanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/10/037, 270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552, 317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488, 725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pL_genes Version 1.0
; SEQ ID NO 131
; LENGTH: 1693
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)..(1512)
; US-10-037-270-131

Alignment Scores:
Pred. No.: 5,3e-79 Length: 1693
Score: 698.50 Matches: 145
Percent Similarity: 60.34% Conservative: 68
Best Local Similarity: 41.08% Mismatches: 115
Query Match: 31.02% Indels: 25
Gaps: 6

US-10-075-846-4 (1-431) x US-10-037-270-131 (1-1693)
Oy      10 SerPheLeuLeuLeuThrPheLeuProGlyGlnVal-----LeuLeuArgValAla 26
Db      76 TCGAGATTAGTAATGTCGACAGTACCAAAATCGAGAGTCTGGGCTTCTCTTTCCT 135
Oy      27 LeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnPrometSerProSer--- 45
Db      136 GTGATGATTAACATGCTGCTTGTGACACAGACCAACCAATGACCAACGAAATGTCATAG 195
Oy      46 -----AspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgLeuArg 63
Db      196 GTGAAAGACACACTGGACAGATTCGCTC-----AAAGGATATGACATTCCTTCGCGG 246
Oy      64 ProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 83
Db      247 CCGGACTTCGAGAGCGCCCGCTGCACGCTTGGGATGCGATGATGCTCCGACATGAC 306
Oy      84 SerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTyrPasn 103
Db      307 ATGCTCTCCGAAGTGAATATGATATACACTCACCATGATATTCACAGCATCTTGCAA 366
Oy      104 AspProArgLysSerTyrArgGluTyrProAspAspSerLeuAspLeuLysSerProSerMet 123
Db      367 GACAAAAGGCTTCTTATTCCTGGAATCCCA---CTGAACCTCACCTTACAGCATATGAGGTA 423

```





```

Db      863 CTAGGATTCACGAGCGGTCTTACAATTGACAAACCATCAGCACCCACTCA6GGAGACCTG 92
Oy      324 ProlysalserYrrAllysAlaIleAspllrTriPmelaValylsLeuleuPheval 34
       ||||| :||||| :||||| :||| :|||||
Db      923 CCAAAGATGCCCTTAGTGCAAGAAGCATTAATTTATCTGATGGGTGCTTTGTTTTG 98
Oy      344 PheaIalaleuleuGIurYrAlaIlaIleAsnPheval 356
       ||| ||||| :||||| :||| :|||
Db      983 TTCCTGCCTCTGCTGGAGTATGCTTTGTAAATTATC 1021

RESULT 13
US-10-211-673-11
: Sequence 11, Application US/10211673
: Publication No. US20030013158A1
GENERAL INFORMATION:
APPLICANT: Le Bourdelles, Beatrice
            Whiting, Paul John
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT
                   OF THE GABA-A RECEPTOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSER: Merck & Co., Inc.
STREET: P.O. Box 2000, 126 E. Lincoln Ave.
CITY: Rahway
STATE: NJ
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM Compatible
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/211,673
FILING DATE: 02-Aug-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB95/02323
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Xu, Yang
REGISTRATION NUMBER: 45,243
REFERENCE/DOCKET NUMBER: T1292
TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-594-1307
TELEFAX: 732-594-4720
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1555 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 47...1402
OTHER INFORMATION:
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-211-673-11

Alignment Scores:
Pred. No.:          9,42e-77           length:        1555
Score:             680.50              Matches:         161
Percent Similarity: 50.11%              Conservative:     63
Best Local Similarity: 36.02%            Mismatches:      150
Query Match:       30.22%               Indels:          75
DB:                9                    Gaps:           7

US-10-075-846-4 (1-431) x US-10-211-673-11 (1-1555)
Oy      15 TrpThrleuProGlyGlnValAleuleuArgValAlaleuAlalysGlucIuValIyssSer 34
       ||||| ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

Dd	48	TGAGACGGC-CGGCGCGGCTCTGGCCCGCGCTCGTCCTGTGGGCGCACAGCTCGCC	106
Oy	35	GLYThrLysGLYSerGlnProMetSer-----ProSerAspPhe	47
Dd	107	GGCACACAGAGCGATGATGATGATCGCGGACATACGCGGCTCCAACTGGAGATCTCTCGG	166
Oy	48	LeuAspLysLeuMetGlyArgThrSerGlyIlyrAspAlaArgIleArgProAsnPhelys	67
Dd	167	CTCCCCCAACTGGACGGGCTGATAGCCGGTATACGCCCGCACTTCCGGGCTGGCATTCGGA	226
Oy	68	GlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLys	87
Dd	227	GGCCCCCGCGTGATGTGGCCCTTCCTGCCCTGGAGGCGGCCACCATCGACACATCTCAAG	286
Oy	88	ThrThrMetAspTryArgValAsnValPheLeuArgGlnGlnIrrPasnAspProArgLeu	107
Dd	287	GCCAAACATGAGATACACCATGACGGGTCTTCCTGCACACAGAGCTGGCGGAGCAGCGCTC	346
Oy	108	SerIlyrArgGlyIlyrProAspAspSerIleuAspLeuAspProSerMetLeuAspSerIle	127
Dd	347	TCTCAAC--AACCAACACCAACGAGAGACCTCGGGCGCGGACACCGCTTGTTGGCAACGTG	403
Oy	128	TrLysProAspPhePhePheAlaAsnGluLysGlyAlaAsnPhelysGluValThr	147
Dd	404	TGCGTGGCGGCACACCTTATGTGTAGACCCCAAGTCCGGCTGTTCACAGACTGACGGTG	463
Oy	148	AspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuIlyrSerIleArgLeuThr	167
Dd	464	GAGAAACAGCTCATCTCCGGCTGACGCGGAGCGGGGGTGTACTACAGCATCCGAATCAC	523
Oy	168	LeuIleLeuSerCysLysMetAspLeuLysAsnPhierProMetAspIleGlnThrCysThr	187
Dd	524	TCCACTGGCGCTGGACATGACCTGGCCAAATTCGCCATTCGAGACAGAGAGTGCATG	583
Oy	188	MetGlnLeuGlnSerSerSerIleuLysSerProLeuProSerLeuSerLeuSerVal	207
Dd	584	CTGGAACCTGGAG-----AGCTAC	601
Oy	208	GlyIlyrThrMetLysAspLeuValPheGluThrPleuGlnAspAlaProAlaValGlnVal	227
Dd	602	GGTTACTCATCGAGAGCATCGTGTACTACTAGTGGAGAGCGACAGACATCCACGGG	661
Oy	228	AlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuGlyLysCys	247
Dd	662	CTGGACAAGCTGCAGCTGGCGCAGTTCACCATACACACACTACCGCTTCACACAGGAGCTG	721
Oy	248	ThrLysHisIyTrasnThrGlyLysPheThrCysIleGluValLysPheHisLeuGlnArg	267
Dd	722	ATGAACCTCAAGAGCCGTCGTGCAGTTCGCCAGCGCTCACACTTCCTCACCTGGGAGG	781
Oy	268	GlnMetGlyIlyrIyLeuIleGlnMetIlyIleProSerLeuLeuIleValIleLeuSer	287
Dd	782	AACCGGCGCGTATCATCATCCAAATCCATACATGCCCTCCGTCGTCGTGGCCATTC	841
Oy	288	TrValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThr	307
Dd	842	TGGGCTCTCTCTGGATCAGCCAGCGCGGGTGGCCCGCCAGGATGTCTCTAGGACTCAC	901
Oy	308	ThrValLeuThrPheThrThrGlnSerSerGlySerArgIleAspLeuProLysValSer	327
Dd	902	ACGGGCGCGAGATGACCAACGCTCATGTGTACGTGCCCTCTCCCTCCCAAGGGCATCA	961
Oy	328	TyrValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIleLeu	347
Dd	962	GCCATCAAGAGCACTGGAGCTGCTACTCTTCGATCTGCTGCTTCTGCTGCTTGGCGGCTG	1021
Oy	348	LeuGluIlyrAlaAlaIleAsnPhe-----	355
Dd	1022	GTGGGTCTCGGCTTTGCTCATTTTCAACGCGGACTACAGGAAGACAGAACGCCAAGCTC	1081
Oy	356	--ValSerArgGlnHisLysGlnPhe-----	363
Dd	1082	AAGGTCTCAGAGCGGACGAGATGTGAGCTGACGAAGCCGCAATTTGCTCTTCTTCGCTC	1141

```
QY 364 -----1leargleuargarglnargarglnargleu 375
DB 1142 TCTGTCGCCGGCAGCAGAGAGTGGCATCTCCGCCGACGGCGGTCCGGG- 1200
QY 376 GluGluaspIleIleGlnIuserArpHerYrPheargIlyTYrGlyLeuGlyHisCys 395
DB 1201 GAACCTGAT-----GGGCTCTACAGGTC 1224
QY 396 LeuGlnAlaArgaspGlyIyProMetGluGlySerGlyIleTySerProGlnPro 415
DB 1225 GGTGGCGGTGAGACAGGAGGAGACGAGAGAGAGAGAGGAGGAGG- 1263
QY 416 AlaProLeuLeuargGly 422
DB 1264 AGCCCGCTCAGAGGCGCAGG 1284

RESULT 14
US-10-239-420-13
; Sequence 13, Application US/10239420
; Publication No. US20030096984A1
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
; FILE REFERENCE: 20629P
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US01/09956
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/193,935
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Dermalcentor variabilis
US-10-239-420-13

Alignment Scores:
Pred. No.: 1,266-75 Length: 1150
Score: 670.00 Matches: 130
Percent Similarity: 64.24% Conservatve: 73
Best Local Similarity: 41.14% Mismatches: 91
Query Match: 29.75% Indels: 22
DB: 9 Gaps: 7

US-10-075-846-4 (1-431) x US-10-239-420-13 (1-1150)
QY 46 AspheLeuaspIySleuMetGlyArGfhrSerGlyTyraSPAlaArgIleargPro--- 64
DB 83 GACATTCGAGCAGAGCTCTC-----AAAACTACGATCGAAGGCCCTGCCGAGC 133
QY 65 AsnphelysGlyProProValAsnValThrcysAsnIlePheIleAsnSerPheSer 84
DB 134 AGTCACCTCGCAAAATGCAACTATTGTGTCATGCGAAATTTACATACGAAGTTTGATCA 193
QY 85 ValThrIystrhrThrMetAspTyraGValAsnValPheLeuArgGlnGlnITrasnsp 104
DB 194 ATAAATCTTGGAACATGACTACGACGAACTGACCTTACCTCCGCACTCGCTCGAC 253
QY 105 ProaIgleuSerTyraGlyTyraProaspSerIleuaspLeu---AspProSerMet 123
DB 254 GAGCGGTACCC---AAATCACGCTATCTGTCGCGCTCGACCTTAATGACCAAGCTG 310
QY 124 LeuaspserIleTyraPlyProaspLeuPheheAlaAsnGlyTyraGlyAlaAsnphHis 143
DB 311 GTACAAATGATATGAGAGCCAGAAAGTTTCTTTCGAAAGCGGAAACAGCGGAGTTCCAA 370
QY 144 GluValThrThrAspAsnIySleuLeuArgIlePheIySasnGlyAsnValLeuTySer 163
```

```
DB 371 TATGTACTGTACTTAACGTCCTGTTAGATCAACCCGAGTGAATATCTTACATG 430
QY 164 IleargleuThrIleLeuSerCysIleuMetAspLeuIyAsnphProMetAspIle 183
DB 431 TTGCGTTAAACTAGAGTCTCTGCAATGAGACCTGTACCGGTACCCATGATTC 490
QY 184 GlnThrCysThrMetGlnLeuGlySerSerIleuLeuTySerProLeuProSerLeu 203
DB 491 CAAGTCTCAGCATCGAAATGCCCTTTTCC----- 523
QY 204 SerLeuSerValGlyTyThrMetIlyAspLeuValPheGluITrPleuGlnAspAlaPro 223
DB 524 -----AAACACACCGAAGAGCGTCTGTAATGTCCGACAGTCAGCT 568
QY 224 AlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgaspGlyIyAsp 243
DB 569 GTCGT---CTCTTCGATTAACCTCAAGTTGCCCAAGTTTGAATTA---GAAAGGTGAC 622
QY 244 LeuGlyCysCysThrIyHisTyraSnrThrcIyIySpherhrcysIleGlyValIySph 263
DB 623 ACGTCTTATGCAAAAGAAAGTTTCACATAGGAGGAAATACGTTGCTGAAGCCGACTTC 682
QY 264 HisLeuGluArgGlnMetGlyTyTyTyIleuIleGlnMetTyTrIleProSerLeuIle 283
DB 683 TATCGACGCTTCCTCGTTATCACATGCTGACAGCTATCTTCCGACACCTTATC 742
QY 284 ValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGly 303
DB 743 GTGTCATCTATGAGGTGTCATTTGCTGACAGGAGGCAATACCGCGGTGTCACC 802
QY 304 LeuGlyIleThrThrValLeuThrMetThrThrIleSerIySerArgIleSerLeu 323
DB 803 CTGGCGCTAACACAGCGCTCTCACCATCTCAATCAAGAGGTCGGATACAGAAACCTG 862
QY 324 ProIyValSerTyraValIyAlaIleAspIleTrpMetAlaValCysLeuPheVal 343
DB 863 CTCTCCGCTCTGATCATCAAGAGCCATCGATGAGATCGATCTTCTTGTTC 922
QY 344 PheAlaIleLeuLeuGluTyraIleAlaIleAsnphValSerArgIle 359
DB 923 TTTGGGCGCTTCTAGATTCACATCTGTCACATATCTCTGAGGCGG 970

RESULT 15
US-10-239-420-4
; Sequence 4, Application US/10239420
; Publication No. US20030096984A1
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
; FILE REFERENCE: 20629P
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US01/09956
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/193,935
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 4
; LENGTH: 2194
; TYPE: DNA
; ORGANISM: Dermalcentor variabilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)...(1315)
US-10-239-420-4

Alignment Scores:
Pred. No.: 3,166-73 Length: 2194
Score: 655.00 Matches: 158
Percent Similarity: 52.53% Conservatve: 91
```

US-10-075-846-4 (1-431) x US-10-239-420-4 (1-2194)	
Best Local Similarity:	33.33%
Query Match:	29.09%
DB:	9
	Mismatches: 133
	Indels: 94
	Gaps: 14

[illegible]

```

OY      356  ValSerArgGlnHisLysGluPheIleArgLeuArgArgGlnArg-----Arg 372
Db      1073  CTTGCTGCGCAGAA-GCAGATGTCGCCGCTTATGCGGAGCGTCGAGCTTCCCAAGA 1131
OY      373  GlnArg-----LeuGluGluAsp 378
Db      1132  TCTGCTTCTTGTGCGGGAACAGACAAATAATGACCCCCGTCACCGTCATCCGAC 1191
OY      379  Ile-----IleGlnGlu 382
Db      1192  GTCCACCCACGCTGCTTGCGCTTACAGACACCGTCCACAGATGCAACCAATGAGCCG 1251
OY      383  SerArgPheArgPheArg----- 388
Db      1252  GGTGCTGTCCCAATCGGCTTTGTTCTTCAACGCACTACTGCGCCCTATTACTGTCT 1311
OY      389  -----GlyTyrGlyLeu----- 392
Db      1312  CTAGTGTGCCATGGCTTCACGTGCTACAGCTGTCGTCCCAAGTCGACGCCATACGGCGG 1371
OY      393  -----GlyHisCysLeuGlnAlaArgAspGlyGlyProMetGluGlySerGlyIleTyr 410
Db      1372  GAAACGGGTGGCTGCGTACGCCACCAAGGAAACGTCGCGCGATGTGAAGAGACTG--- 1428
OY      411  SerProGlnProProAlaProLeuLeuArgGluGluGluPhe 424
Db      1429  ---CCATTCACCGACGACGCTCTGTGGTGAAGAGAAAGCTTACA 1467

```

Search completed: July 1, 2003, 01:01:34  
Job time : 248.481 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:00:29 ; Search time 2265.88 seconds  
(without alignments)  
3080.588 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTTVPATLFLTLTLTPGQ.....PQPAPLREGETTRKLYVD 431

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-O=/cg22\_1/USPTO\_spool/US10075846/runat\_25062003\_163648\_5157/app\_query.fasta\_1.782  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846.ecgn.1.1.3724.etrnat.25062003.163648.5157 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSBLOCK=100 -LONGIOS  
-REV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estlro:\*  
8: em\_hlc:\*  
9: gb\_estcl:\*  
10: gb\_estcl2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_ggs:\*  
18: gb\_ggs\_hum:\*  
19: em\_ggs\_inv:\*  
20: em\_ggs\_pln:\*  
21: em\_ggs\_vrt:\*  
22: em\_ggs\_fun:\*  
23: em\_ggs\_mam:\*  
24: em\_ggs\_mus:\*  
25: em\_ggs\_other:\*  
26: em\_ggs\_pro:\*  
27: em\_ggs\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1033.5	45.9	662	9	AUI69868
2	877.5	39.0	779	14	BQ443713
3	854.5	37.9	678	10	AV729257
4	853.5	37.9	2991	11	BC027094
5	832.5	37.0	2593	11	BC022502
6	789.5	35.1	692	12	BE981841
7	772.5	34.3	720	12	BE981459
8	746.5	33.1	817	12	BG404477
9	672.5	29.9	975	14	BQ938794
10	631.5	28.0	2475	11	AK013727
11	631	28.0	893	9	AL538200
12	619.5	27.5	802	14	BQ042344
13	611	27.1	817	9	AUI33323
14	605	26.9	1942	11	BC029850
15	570.5	25.3	948	9	AL539373
16	555	24.6	3257	11	BC026337
17	548.5	24.4	536	12	BF906462
18	548	24.3	552	13	BM440664
19	545	24.2	957	9	AL572853
20	519	23.0	586	10	AM122941
21	510	22.6	594	12	BF937654
22	504	22.4	2708	11	AK018768
23	501	22.2	690	14	BQ044592
24	499	22.2	723	13	B1736340
25	499	22.2	908	13	B1825190
26	498.5	22.1	663	10	BB653397
27	497	22.1	831	13	B1663922
28	488.5	21.7	1134	14	BQ667322
29	488	21.7	1068	13	BM547796
30	482.5	21.4	623	10	AM077068
31	479	21.3	1078	13	B1757807
32	470	20.9	785	14	BM964113
33	469.5	20.8	843	13	B1913857
34	465	20.6	1047	14	BQ067799
35	461.5	20.5	501	10	AM280976
36	460.5	20.4	912	13	B1829118
37	460.5	20.4	975	13	B1553703
38	460	20.4	867	12	BG189196
39	460	20.4	918	13	B1824761
40	459.5	20.4	582	13	B1681394
41	457	20.3	567	13	BM129587
42	457	20.3	570	13	BM129116
43	456	20.2	814	12	BG706397
44	446.5	19.8	861	13	B1535004
45	445.5	19.8	700	12	BG404089

## ALIGNMENTS

RESULT 1  
AUI69868  
LOCUS AUI69868 662 bp mRNA linear EST 29-JAN-2001  
DEFINITION AUI69868 Ol-br-ad cDNA *Oryzias latipes* cDNA clone br5332, mRNA  
ACCESSION AUI69868  
VERSION AUI69868.1 GI:12591937  
KEYWORDS EST.  
SOURCE Japanese medaka.  
ORGANISM *Oryzias latipes*  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Aluettinomorpha;  
Belontiiformes; Adrianichthyidae; Oryziinae; *Oryzias*.  
REFERENCE 1 (bases 1 to 662)

AUTHORS Mita,K., Ishikawa,Y. and Yamauchi,M.  
 TITLE Establishment of cDNA database of medaka, *Oryzias latipes*  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Mita K

Genome Research Group  
 National Institute of Radiological Sciences  
 Anagawa 4-9-1, Inage, Chiba 263-8555, Japan  
 Email: kmite@irs.go.jp  
 method: uni-directional sequence direction: sequenced from T3 primer  
 (5' -> 3').

# FEATURES

source Location/Qualifiers  
 1. 662  
 /organism="Oryzias latipes"  
 /strain="HN1"  
 /db\_xref="taxon:8090"  
 /clone="pr5332"  
 /clone\_1lb="OI-br-ad cDNA"  
 /sex="female/male mixed"  
 /tissue\_type="brain"  
 /dev\_stage="adult"  
 BASE COUNT 155 a 205 c 164 g 138 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,66e-121 Length: 662  
 Score: 1033.50 Matches: 193  
 Percent Similarity: 88.09% Conservative: 14  
 Best Local Similarity: 82.13% Mismatches: 13  
 Query Match: 45.89% Indels: 15  
 DB: 9 Gaps: 2

US-10-075-846-4 (1-431) x AU169868 (1-662)

QY 66 PhelysG1yProProValAsnValThrycAsnIlePheIleXnsrPheSerVal 85  
 |||||  
 Db 2 TTTAAAGGTCACCTTAACGCTGACCTGACCATTTTATCAGACGCTTTGGATCATT 61  
 QY 86 ThrlYsThThMetAspTyrArGValAsnValPheLeuArGInGIntPrAsnAPro 105  
 : : : : :  
 Db 62 GGTGAACACCATGAGCTACAGAGTGAACATCTTCTGAGCAGCAGTGAACGACCC 121  
 QY 106 ArgLeuSerTyrArGInuTyrProAspAspSerLeuAspLeuAspProSerMetLeuAsp 125  
 |||||  
 Db 122 CGCTGGCTGACGACGATATCCGACGACTCGCTGATCTGGACCCGTCATGTTGGAC 181  
 QY 126 SerIleTPrLySProAspLeuPhePheAlaAsnGluValAsnGlyAlaAsnPheHisGluVal 145  
 |||||  
 Db 182 TCCATGTGAGAGCGGATCTGTTCTTGTATAGAGAAGGGGCCAATCTCCGAGGTC 241  
 QY 146 ThrlThAspAsnLysLeuLeuArGIllePheLysAsnGlyAsnValLeuTyrSerIleArG 165  
 |||||  
 Db 242 ACCACGACACACAGAGCTGCGCATCTCCAAAATGGCAGACGCTGTACACATACGAC 301  
 QY 166 LeuThrlLeuIleuSerLysLeuMetAspLeuLysAsnPheProMetAspIleGIntPr 185  
 |||||  
 Db 302 ATACCTCTATCTGCGCTGCGCCCATGATCTGMAAAGTCTCCCATGATGTGCAGACC 361  
 QY 186 CysThrMetGlnLeuLysSerSerIleLeuLysSerProLeuProSerLeuSerLeu 205  
 |||||  
 Db 362 TGCATCATGACAGTGAAG----- 379  
 QY 206 SerValGlyTyrThrMetLysAspLeuValPheGluTPrLeuGlnAspAlaProAlaVal 225  
 |||||  
 Db 380 ACCTTCGGCTACACCATGAGACGACCTCATCTTCGACTGG--GACGAGAAGGGCGCGTG 436  
 QY 226 GlnValAlaGluGlyLeuThrlLeuProGlnPheIleLeuArGAspLysAspLeuGly 245  
 |||||  
 Db 437 CAGGTGGCGGACGCGCTGACGCTGCTCATCTCCTAAGAGAGAGAGAGAGAGAGAGAG 496  
 QY 246 CysCysThrlYsThrlYsArnThrlGlyLysPheThrCysAlleGluValLysPheHisLeu 265  
 |||||  
 Db 497 TACTGCACCAAGACACTACACACAGGTAATTCACCTGATGAGGCTTCACGCTG 556

QY 266 GluArGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIle 285  
 |||||  
 Db 557 GAGCGACACATGGGCTACTACTGATCCGAGATGATCATCCCTGCTGCTCATC 616

QY 286 LeuSerTPrValSerPheTrpIleAsnMetAspAlaAlaProAla 300  
 |||||  
 Db 617 CTGTCTGGGTCTCTCTGTGATCAACATGAGCGCGCGCGCC 661

## RESULT 2

BO443713 779 bp mRNA linear EST 29-MAY-2002  
 LOCUS UI-M-EMO-bxh-j-02-0-UI.r1 NIH\_BMAP-EMO Mus musculus cDNA clone  
 DEFINITION IMAGE:5708257 5', mRNA sequence.  
 ACCESSION BO443713  
 VERSION BO443713.1 GI:21246825  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1 (bases 1 to 779)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgrabbs@email.nih.gov  
 Tissue Procurement: Dr. James Lin, University of Iowa

cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)  
 Seq primer: ptx-5.

## FEATURES

source Location/Qualifiers  
 1. 779  
 /organism="Mus musculus"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5708257"  
 /clone\_1lb="NIH\_BMAP-EMO"  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 15.5 dpc"  
 /lab\_host="DHIOB (T1 phage resistant)"  
 /note="Organ: brain; Vector: ptx-asc; Site\_1: EcoR I;  
 Site\_2: Not I; The library was constructed according to  
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured mRNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with an  
 oligo-dT primer containing a Not I site. Double stranded  
 cDNA was size selected according to mRNA size fraction,  
 ligated with EcoR I adaptor, digested with Not I, and then  
 cloned directionally into ptx-asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GTGGCTGGAA. This library was created for the  
 University of Iowa Mouse Brain Molecular Anatomy Project  
 (BMAP). 'Gene Discovery in the Developing Mouse Nervous  
 System', supported by National Institutes of Mental Health  
 (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 222 a 167 c 165 g 225 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.52e-101 Length: 779  
 Score: 877.50 Matches: 174  
 Percent Similarity: 78.99% Conservative: 14  
 Best Local Similarity: 73.11% Mismatches: 29  
 Query Match: 38.97% Indels: 21  
 DB: 14 Gaps: 4

US-10-075-846-4 (1-431) x BQ443713 (1-779)

```
Oy 1 MetThrLeuValProAlaThrLeuSerPheLeuLeuTyrPThrLeuProGlyIn 20
Db 126 CTAGTGAACATTGTTGACAGCCTTGTGCTTTCTTTCTTA-----GGGACA 170
Oy 21 ValLeuLeuArgValAlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySer 39
Db 171 AACCACTTCAGAGGAAGCTTTCTGCAAAAGACATGACCTCAGGTCTGGAAGAACATCCCTCG 230
Oy 40 GlnProMetSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAsp 59
Db 221 CAGACCCCTATCTCTCCATGATTTCTTGGATTAATTAATGCAAGAACATCAGGATATCAT 290
Oy 60 AlaArgLeuLeuProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIle 79
Db 291 GCAGAAATCAGGCAAAATTTTAAAGCTCCAGTAAACGTTACTTGCAGATTTTTCATC 350
Oy 80 AsnSerPheSerSerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArg 99
Db 351 AACAGTTTGGATGATCAGACAGAAACACCATGACCTACCGAGTGAACATTTTCTGAGA 410
Oy 100 GlnGlnTTPAsnAspProArgLeuSerTyrArgGlyTyrProAspAspSerLeuAspLeu 119
Db 411 CAGCAGTGGAAATGATTCAGGCTGGCAGATGAGTACCCAGATGATTCCTGGGATTTG 470
Oy 120 AspProSerMetLeuAspSerIleTyrLysProAspLeuPheAlaAsnGlyLysGly 139
Db 471 CATCCCTCAATGTTGGATTCATTTTGGAAACCGGATTTGTTTCCCAATCAAAAGGCT 530
Oy 140 AlaAsnPheHisGluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159
Db 531 GCCAATTTTCATGATGATGCACCACTGACAAACAAATGTTGGCGATTTCCAAATATGGCAAA 590
Oy 160 ValLeuTyrSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnIle 179
Db 551 GTGCTCTACAGATTAGACTCACCCTTGAATTTATCTTGTCCATGAGCTGAGAAAGCTTT 650
Oy 180 PrometAspIleGlnThrCysThrMetGlnLeuGlnSerSerIleLeuLysSerPro 199
Db 651 CCAATGGATGTCACCACTGATCATGACGCTGAG----- 686
Oy 200 LeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrLeu 219
Db 687 -----AGTTTGGGTACACCATGAAATGACCTGATATTTGATGCTGTA 728
Oy 220 GluAspAlaProAlaValAlaGluGlyLeuThrLeuProGlnPheIle 237
Db 729 AGTGATGTGCA---GTACCAAGTGTCTGAAAGACTCACCTGCCCCAGTTTATT 779

RESULT 3
AV729257 678 bp mRNA linear EST 17-OCT-2000
LOCUS AV729257 HTC Homo sapiens cDNA clone HTCHE03 5', mRNA sequence.
DEFINITION AV729257
ACCESSION AV729257.1 GI:10838678
VERSION AV729257.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 678)
AUTHORS Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X.,
Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu,
S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,
Chen, J., Chen, Z. and Han, Z.
Homo sapiens cDNA clones
Unpublished (2000)
CONTACT: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203 P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
```

```
Email: hanzg@hgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HTCHE03"
/clone_lib="HTC"
/tissue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLK"
/notes="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 199 a 148 c 150 g 180 t I' others
ORIGIN
Alignment Scores:
Pred. No.: 1.74e-98 Length: 678
Score: 854.50 Matches: 173
Percent Similarity: 81.03% Conservative: 15
Best Local Similarity: 74.57% Mismatches: 21
Query Match: 37.94% Indels: 23
DB: 10 Gaps: 4
US-10-075-846-4 (1-431).x AV729257 (1-678)
Oy 26 AlaLeuAlaLysGlu---GluValLysSerGlyThrLysGlySerGlnProMetSerPro 44
Db 38 GCTTTCGCAAGACCATGATGCTCCAGGTCTGGAAGAACACCTTCACAGACCTATCTCT 97
Oy 45 SerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgIleArgPro 64
Db 98 TCAGATTTCTTGACAGATTAAATGGAAGACATCAGATATGATGCAAGATTCAGGCCA 157
Oy 65 AsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSer 84
Db 158 AATTTTAAAGCTCTCCAGTAAACGTTACTTGCATTAATTTTATCAACGTTTGGATCA 217
Oy 85 ValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnIleTyrAsnAsp 104
Db 218 GTACAGAAACGACCATGACCTACCGAGTGAATATTTTCTAGACACACAGTGAATGAT 277
Oy 105 ProArgLeuSerTyrArgGlyTyrProAspAspSerLeuAspLeuAspProSerMetLeu 124
Db 278 TCACGGCTGGCTGACAGTACGATGACCATGATGATTTTCTAGACACACATGCTA 337
Oy 125 AspSerIleTyrLysProAspLeuPhePheAlaAsnGlyLysGlyAlaAsnPheHisGlu 144
Db 338 GACTCCATTTGGAAACCAAGATTTGTTCTTGGCAATGAGAAAGGTCCTCAACTTCACGAT 397
Oy 145 ValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuTyrSerIle 164
Db 398 GTACACCATGACCAACAAATTTGCTACGATTTTGGAAAAATGGCAAGTCTTACAGTATC 457
Oy 165 ArgLeuThrLeuIleLeuSerCysLeuMetLysAspLeuLysAsnPheProMetAspIleGln 184
Db 458 AGACATCACCTGACCTTATCTGCTCCATGAGCACTTGAAACACTTCCAGTATGATGTCAG 517
Oy 185 ThrCysThrMetGlnLeuLeuSerSerIleLeuCysSerProLeuProSerLeuSer 204
Db 518 ACCCTGACATGACAGCTGGAG----- 538
Oy 205 LeuSerValGlyTyrThrMetLysAspLeuValPheGlnTyrLeuGlnLysAspAlaProAla 224
Db 539 ---AGTTTGGGTACACCATGATGATGACCTGATTTGAGTGGTAAAGTATGATGCTCA--- 592
Oy 225 ValGlnValAlaGluGlyLeuThr--LeuProGlnPheIleLeuArgAspGlu--LysAspL 244
Db 593 GTGCAAGTGTCTGAAAGATTTGACCTTGGCCCAAGCTTATTTTGAAGAAAGACAGCGAGCC 652
Oy 244 euGlyCysCysThrLysHisTyrAsnThrGly 254
Db 653 TTGGC-----TACTTGACTGGA 669
```





DEFINITION Homo sapiens, glycine receptor, beta, clone IMAGE:4792516, mRNA.  
 ACCESSION BC022502  
 VERSION BC022502.1 GI:18490294  
 KEYWORDS HTC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2593)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
 CONTACT: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Miklos Palokovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www.shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcdepaxil.stanford.edu](mailto:mcdepaxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 32 Row: b Column: 19  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA q1: 4504022  
 This clone has the following problem: frame shifted.  
 FEATURES  
 source  
 1..2593  
 /organism="Homo sapiens"  
 /db\_xref="LocusID:2743"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4792516"  
 /tissue\_type="Brain, hypothalamus"  
 /clone\_id="NIH\_MGC\_96"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"  
 BASE COUNT 824 a 466 c 488 g 815 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.54e-94 Length: 2593  
 Score: 832.50 Matches: 186  
 Percent Similarity: 63.78% Conservative: 64  
 Best Local Similarity: 47.45% Mismatches: 89  
 Query Match: 36.97% Indels: 54  
 DB: 11 Gaps: 11  
 US-10-075-846-4 (1-431) x BC022502 (1-2593)  
 Oy 8 ThrleuSerPheleuLeuLeuTrPThleuProGlyInValleuLeuAArgValAlaleu 27  
 Db 116 ACAACCTCCCTTTAAATTTAATTTCTCTG-----TAAGTGAAGAAGCCATAT 163  
 Oy 28 AlAlaySgluValIySergIyThrlYsglySergInPromet----- 42  
 Db 164 TCTAAGCAAAAGCTTCAAGAAAGGAGGAAAAAGAACGACTATCTATGCCCATCT 223  
 Oy 43 -----SerProSerAspPheLeuAsp 49  
 Db 224 CAGCAGTCAGCAGAGACCTTCCCGAGTACTGCCAACCTCCACATGATATCTTGAC 283  
 Oy 50 LysleuMetGlyArgThrSergIyTyraSPalaArgleArgProAsnPhelYsglyPro 69  
 Db 284 AGGTATTG-----GTCAATTATGATCCAGATTAAGCAAACTTAAAGGCAATT 334

Oy 70 ProValAsnValThrCysAsnIlePheIleAsnSerPheSerSerValThrlYstThr 89  
 Db 335 CTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 394  
 Oy 90 MetAspTyraArgValAsnValPheLeuArgGlnIleTrpAsnAspProArgLeu 107  
 Db 395 ATGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454  
 Oy 108 -----SerTyraArgGlyTyraProAspSerIleuAspSerIleuAspSerIleuAsp 125  
 Db 455 CCCAGTATTTTACG-----GTTCAATGCTGACAGTGCAGTGCAGTGCAGTGCAGT 508  
 Oy 126 SerIleTrpIysProAspPhePhePheAlaAsnGluYsglyAlaAsnPhelSglVal 145  
 Db 509 TGTTATGAAACCTGATTTATTTTGGCAATGAAAGAGTGCATTTTCATGATGATG 568  
 Oy 146 ThrThraSPasnIySleuLeuArgIlePheIySasnIySasnValLeuTyraSerIleArg 165  
 Db 569 ACCCAGGAAACATCCCTCTTATTTATTTCTGATGAGATGCTTCTTCAGCATGAGG 628  
 Oy 166 LeuThrIleuLeuSerCysIleuMetAspIleuYsAsnPhelProMetAspIleGlnThr 185  
 Db 629 TTATCTATTTACTTTTATGATGCTTGGACTTGACATTTTCCAAATGATACACACGT 688  
 Oy 186 CysThrMetGlnLeuGlnSerSerSerIleuYsSerProleuProSerIleuSerLeu 205  
 Db 689 TGCAAGATGCACGACG----- 706  
 Oy 206 SerValGlyTyrrThrMetIySAspLeuValPheGlnTrPleuGlnAspAlaProAlaVal 225  
 Db 707 AGCTTGTATACACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 763  
 Oy 226 GlnValAlaGlnGlyLeuThrlleuProGlnPhe-----IleLeuArgAspGluYsAspLeu 244  
 Db 764 CAATTA--GAAAAAATGCTGCTCCCAATTTGATATCAAAAAGAGATATTGATAT 820  
 Oy 245 GlnYsCysThrIySHISlyr---AsnThrlYsPheThrlYsIleGlnValIySAsp 263  
 Db 821 GGTAACTGTACAAAATACATAAAGCAGCGGCTACACATCGTGGAAGTATCTTC 880  
 Oy 264 HisLeuGlnArgGlnMetGlyTyrrIleuIleGlnMetTyrrIleProSerLeuIle 283  
 Db 881 ACCCTGAGAGGCGAGTGGCTTTACATGATGAGGCTTCACGCCCAACCTCTCTCAT 940  
 Oy 284 ValIleuSerTrpValSerPheTrpIleAsnMetSPAlaAlaProAlaArgValGly 303  
 Db 941 GTTGTCTCTCTGCTTCTCTCTGATCAACCCGAGCGGAGTGTGCTGCGAGTGC- 999  
 Oy 304 LeuGlyIlePThrThrValIleuThrMetThrThrlGlnSerSergIySerArgAlaSerLeu 323  
 Db 1000 CTGGGTATCTTCTTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCT 1059  
 Oy 324 ProIySValSerTyraValIySAlaIleAspIleTrpMetAlaValCysLeuLeuPheVal 343  
 Db 1060 CCCAAAGTTTCTTCTGAGAGCTTGTATGATTTGCTTATGCTTCTCTCTCTGCTTCTG 1119  
 Oy 344 PheAlaAlaLeuLeuGlnTyraAlaAlaIle-----AsnPhelValSerArgGln 359  
 Db 1120 TTTTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1179  
 Oy 360 HisIySgluPheIleArgLeuArgArgGlnArg 371  
 Db 1180 GAAGCTGAAAGGACGACATATGCTAAGCTGACCA 1215  
 RESULT 6  
 BE981841 692 bp mRNA linear EST 29-APR-2002  
 LOCUS UT-M-GG0P-bdd-h-07-0-UT.s1 NIH\_BMAP.Ret4.s2 Mus musculus cDNA clone  
 DEFINITION UT-M-GG0P-bdd-h-07-0-UT 3', mRNA sequence.  
 ACCESSION BE981841  
 VERSION BE981841.1 GI:10651376  
 KEYWORDS EST.

SOURCE mouse mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 692)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 9704447  
COMMENT  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: MEST@mail.nih.gov  
Oligo-dT track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA library Preparation: M.B. Soares lab clone distribution: Researchers may obtain BMAP CDNA clones from RESEARCH GENERICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP CDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine  
Seq primer: M13 Forward  
POLYA-No.

BASE COUNT	190 a	159 c	144 g	198 t	1 others
ORIGIN					

Alignment Scores:	
Pred. No.:	4,08e-90
Score:	789.50
Percent Similarity:	78.95%
Best Local Similarity:	75.12%
Query Match:	35,06%
DB:	12
Length:	692
Matches:	159
Conservative:	8
Mismatches:	28
Indels:	16
Gaps:	2

US-10-075-846-4 (1-431) x BE981841 (1-692)

QY	8	ThIeuSerPheIeuLeuIeTrp--ThIeuProGluValIeuAucValI	26
Db	110	ACACCTTCCGACACACTGTTCTGGATTTTACTTCTGGGAACGCCACGTTACTGAGTTT	165
QY	26	aIeuAlaIysGluGluValIysSerGlyThrIysGlySerGlnProMetSerProSerAs	46
Db	170	GGTGGCCACAAAGAAACAAACAGTGCCTGATCTCGAAGTGTCTCCAAATGACCTTCGA	225
QY	46	pPheIeuAspIysIeuMetGlyArgThrSerGlyTyrIspAlaargIleargProAsnPh	66
Db	230	TTTTTCGGACAACTAATGGGGGACACATCGGGGTATGTGCAAAATCAACACCAACTT	285
QY	66	eIysGlyProProValAsnValIthCysAsnIlePheIleAsnSerPheSerSerValIth	86

Db	290	CAAGGCTCCACAGTTAATGTCACATGCACATATTCAATTAACACCTTTGGCTCCATTGTC	349
Oy	86	rlYsThrThrMetCAspTyraValAsnValaPheLeuArgGlnIleTTPAsnAspProAr	106
		.....	
Db	350	AGAGCGCACTATGGATTACAGAGTAAACATTTTTCTTGTCAGAGTGAATGATCCCTGC	409
Oy	106	gluSerSerIyArgGluTyTProAspAspSerLeuAspLeuAspProSerMetLeuAspSe	126
		.....	
Db	410	TCCTTGCAATACAGTGAATATCCCTGACGATTCAATTAACCTTGACCCGCTCATATGTGGATTG	469
Oy	126	rIleTrpLysProAspLeuPhePheAlaAsnGluIySgIyAlaAsnPheHisGluValaIrh	146
		.....	
Db	470	CATTGGAAACCTGCCTTGCTTGTGTAATGAGAAGGGCGCTTACTTCCACAGAGTGC	529
Oy	146	rThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValaLeuTySerSerIleArgLe	166
		.....	
Db	530	CACAGATTAACAAGCGCTTAAGAAATTTTANANAATGTGAATGTCCTTTATTCAATTAAGGTT	589
Oy	166	uThrLeuIleLeuSerCysLeuMetAspLeuLysAsnIlePhePromeTAspIleGluThrCy	186
		.....	
Db	590	GACATTTAACACTCTCTCTGCTCCATATGGAATCTCAAGAAATTTCCCAATGGATGTCAACAACATG	649
Oy	186	sThrMetGlnLeuGlnSerSerSerIleLeuCysSerProLeuProSerLeuSerLeuSe	206
		.....	
Db	650	CATAATGCAACTTGAA-----AG	667
Oy	206	rValGlyTyTThrMetLysAspLeu	214
		.....	
Db	668	CTTTGGGTACACAATGAATGATCTCT	692

RESULT 7	
BE981459	
LOCUS	BE981459 720 bp mRNA linear EST 29-APR-2001
DEFINITION	UI-M-CG0P-bdc-d-08-0-UI.s1 N1H_BMAP_Ret4_S2 Mus musculus cDNA clone.
ACCESSION	UI-M-CG0P-bdc-d-08-0-UI 3', mRNA sequence.
VERSION	BE981459
KEYWORDS	BE981459.1 GI:10650587
SOURCE	EST.
ORGANISM	house mouse. Mus musculus

REFERENCE  
AUTHORS  
TITLE

Euarctota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 720)

Ronald M.F., Lennon G. and Soares M.B.

Normalization and subtraction: two approaches to facilitate gene

JOURNAL  
MEDLINE  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
Contact: Chin, H

COMMENT

Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: MEST@mail.nih.gov

Oligo- or track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMP4 cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine

Seq primer: M13 Forward

POLYA-No.

```

DEFINITION Homo sapiens, glycine receptor, beta, clone IMAGE:4792516, mRNA.
ACCESSION BC022502
VERSION BC022502.1 GI:18490294
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2593)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT MGC help desk
Contact: gcapbs-remail.nih.gov
Email: gcapbs-remail.nih.gov
Tissue Procurement: Miklos Pavkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 32 Row: b Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504022
This clone has the following problem: frame shifted.
FEATURES
Source location/qualifiers
1..2593
/organism="Homo sapiens"
/db_xref="LOCUSID:2743"
/db_xref="Laxon:9606"
/clone="IMAGE:4792516"
/tissue_type="Brain, hypothalamus"
/clone_lib="NH_MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
BASE COUNT 824 a 466 c 488 g 815 t
ORIGIN
Alignment Scores:
Pred. No.: 1.54e-94 Length: 2593
Score: 832.50 Matches: 186
Best Local Similarity: 63.78% Conservative: 64
Percent Similarity: 47.45% Mismatches: 89
Query Match: 36.97% Indels: 54
DB: 11 Gaps: 11
US-10-075-846-4 (1-431) x BC022502 (1-2593)
Oy 8 ThrLeuSerPheLeuLeuLeuTriphrLeuProGlyGlnValLeuLeuArgValAlaLeu 27
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 ACAATGCCCTTTTAAATTTTAAATTTCCCTTG-----TAAGTGAGAAAGCCTAT 163
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 28 AAlayGcUGlnUVallySerGlyThrlyrGcGlySerGlnPromet----- 42
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164 TCTAAGGAAAAGCTTCAAGAAAGGGAAGGGAAGAAAGAACATATATGCCCATCT 223
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 43 -----SerProSerAspPheLeuAsp 49
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 224 CAGCAGTCAGCAGACGACCTTGCCCGATACCTGCCAACATCCACATGCAATATCTTGAAC 283
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Oy 50 LysLeuMetGlyArgThrSerGlyrAspAlaArgIleArgProAsnPheLysGlyPro 69
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 284 AGGTAATATG-----GTCAAGTTATATCCACAGATPAAGACCAAACTTCAAGGCAATT 334
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

OY	70	ProValAsnValIthrCysAsnIlePheIleAsnSerPheSerValIthrIysThrThr	89
Db	335	CCCTGTTGATGACAGTACCAACATTTTATTAACAGTTTGGATCCATTTCAAGAAACACA	394
OY	90	MetAspTyrArgValAsnValPheLeuArgIleInTPasAspProArgLeu	107
Db	395	ATGCACTATAGATTACATCTTCTGAGACAAAATGGAATGACCCAGGTAAAGCTC	454
OY	108	-----SerTyrArgIleTyrProAspSerLeuAspProSerMetLeuAsp	125
Db	455	CCCAAGTATTTTAGG-----GGTTCAAGATGCACAGAGTGATCCAAACATGTACAAAG	508
OY	126	SerIleTyrPyrProAspLeuPhePheAlaAsnIleIuysGlyAlaAsnPheHisGluVal	145
Db	509	TGTTTATGAAACCTGATTTATTTTTCGCAATATAAAAGCTCCAAATTTTCATGATGCTG	568
OY	146	ThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyAsnValIleTyrSerIleArg	165
Db	569	ACCCAGGAAAAACATCCTCTCTTATATTTTGGTGATGAGATGTCTTTGTCAGCATGAGC	628
OY	166	LeuThrIleuIleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGluThr	185
Db	629	TTATCTATTACTCTTCATGCCCTTTGGACTTGACATGTTTCCAAATGATACACACAGT	688
OY	186	CysThrMetGluLeuGluSerSerIleLeuCysSerProLeuProSerLeuSerIleu	205
Db	689	TGCAAGAGACACCTGAC-----	706
OY	206	SerValGlyTyrThrMetLysAspLeuValPheGluTyrPleuGluAspAlaProAlaVal	225
Db	707	AGCTTTGGTTACACACATGATGATTTACGATTTATCTGCGAGTCAGAGATCCT---CTG	763
OY	226	GluValAlaIleGluGlyLeuThrLeuProGluPhe---IleLeuArgAspGluLysAspLeu	244
Db	764	CAATTA---GAAAAAATGTCCTTGCCTCAATTTGATCAAAAAGAAAGATTTGAATAT	820
OY	245	GlyCysCysThrLysHisIstYr---AsnThrGlyLysPheThrCysIleGluValLysPhe	263
Db	821	GGTAACTATCAAAATACTATAAAGCGCAGGGCTACTACACATGCGTGAAGTCAATCTTC	880
OY	264	HisLeuGluArgIleMetGlyTyrTyrIleuIleGluMetTyrIleProSerLeuLeuIle	283
Db	881	ACCTGAGGAGGACGAGCGCTTTTACATGATGGGGGCTACGCCCAACCTGCTCATTT	940
OY	284	ValIleLeuSerTyrPvalSerPheTyrIleAsnMetAspAlaAlaProAlaArgValGly	303
Db	941	GTTGTTCTCTCTGCTCTTCTCTTCGATCAACCGGACGGAGTGCACGAGTGC-	999
OY	304	LeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeu	323
Db	1000	CTGGGTATCTTCTCAGTCTCTCAGCTGGCCCTCTGAGTGCACACACCTTGGCCGTAGCTT	1055
OY	324	ProLysValSerTyrValLysAlaIleAspIleTyrMetAlaValCysLeuLeuPheVal	343
Db	1060	CCCAAGATTTCCTATGTAAGGCTCTGATTTGGCTTATTTGCTTCCTCTCTTTGGG	1119
OY	344	PheAlaAlaLeuLeuGluTyrAlaAlaIle-----AsnPheValSerArgGln	359
Db	1120	TTTGGTTCCTCCGTCGAGTATGCAATTTGCCAGGTATGCTGAACACCCCAAAAGGCTT	1179
OY	360	HisLysGluPheIleArgLeuArgTyrArgIleArg 371	
Db	1180	CAAGCTGAAAAAGCCAGAAATTCCTAAGCTGACGAA 1215	
RESULT 6			
BE981841			
LOCUS			
DEFINITION			
UI-M-CG0P-bdd-h-07-0-UI.s1 NIH BMP Ret4.S2 Mus musculus cDNA clone			
ACCESSION			
BE981841			
VERSION			
BE981841.1			
KEYWORDS			
EST.			

SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE Euxariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 692)  
TITLE Ronaldo,M.F., Lennon,G. and Soares,M.B.  
JOURNAL Normalization and subtraction: two approaches to facilitate gene  
MEDLINE discovery  
COMMENT Genome Res. 6 (9), 791-806 (1996)  
9704447  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: m5t@mail.nih.gov  
Oligo-dT track not found, Not I site shown in beginning of sequence  
is likely internal to the message. cDNA Library Preparation: M.B.  
Soares lab clone distribution: Researchers may obtain BMAP.cDNA  
clones from RESEARCH GENERICS. It should be noted that Bento Soares  
is generating a small number of additional specialized  
non-redundant arrays of BMAP cDNAs whose availability will be  
considered under appropriate and limited collaborative arrangements.  
The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
Yale University School of Medicine  
Send primer: M13 Forward  
POLYA-No.

BASE COUNT	190 a	159 c	144 g	198 t	1 others
ORIGIN					

Alignment Scores:	
Pred. No.:	4,08e-90
Score:	789.50
Percent Similarity:	78.95%
Best Local Similarity:	75.12%
Query Match:	35,06%
DB:	12
Length:	6932
Matches:	157
Conservative:	8
Mismatches:	28
Indels:	16
Gaps:	2

US-10-075-846-4 (1-431) x BE981841 (1-692)

[illegible]

Db	290	CAAAAGTCTCCAGTTAATGTCACTGCACACATATTTCATTAACACAGCTTTGGCCATTCGC	349
QY	86	rlYstrThrthMetAspTyrAlaValAsnValPheLeuArgInlTrpAsnAspProAr	106
Db	350	AGAGACGCACTATGAGATTACAGAGTAACATTTTTCTTCCTCCAGAGTGGAAATGATCCTCG	409
QY	106	gluSerIyrlArguIuTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSe	126
Db	410	TCTTCATACAGCTGAATATCCTGCAGCATTCAATTAGACCTTGACCCGCTTATGTTGGATTTC	469
QY	126	rlleTrpIyPsrProAspLeuPhePheAlaAsnIuIySgIyAlaAsnPheHisGluValIrh	146
Db	470	CATTGGAAACCTGCAGCTTGTCTTGGCTAATGAGAAGGGGGGCTAACTTCACAGAAAGTCAC	529
QY	146	rThrsAspAsnIySleuLeuArgIlePheIyAsnSngIyAsnValLeuTyrSerIleArgIle	166
Db	530	CACACATATAACAGCGCTAAGAAATTTTCANAAATGGAAATGTCCTTATTATTCATATAGGTTT	589
QY	166	uThrLeuIleLeuSerCysLeuMetAspLeuIyAsnIlePhePrometAspIleGlnTrpCty	186
Db	590	GACATTAAACACTCTCTCTGTCACAAATGGATTCACAAGATTTCCCAATGGATGTACAAACATG	649
QY	186	sThrMetCInleuIySerSerIleLeuCysSerProIeuProSerLeuSerLeuSe	206
Db	650	CATATACCACTTGA-----AG	667
QY	206	rValGlyTyrTrpMetIyAspLeu	214
Db	668	CTTTGGGTACACATGAATGATCTTC	692

RESULT 7	
BE981459	
LOCUS	BE981459 720 bp mRNA
DEFINITION	UI-M-CG0P-bdc-d-08-0-UI.s1 N1H_BMAP.Ret4_S2 Mus musculus cDNA clone.
ACCESSION	BE981459
VERSION	BE981459.1 GI:10650587
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE  
AUTHORS  
TITLE

Emmariyola: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
Mullaria: Eutheria: Rodentia: Sciurognathi: Muridae: Mus  
1 (bases 1 to 720)  
Bonaldio,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene

JOURNAL	Genome Res. 6 (9), 791-806 (1996)
MEDLINE	97044477
COMMENT	Contact: Chin, H

Tel: 301 443 1706  
Fax: 301 443 9890

Email: MEST@mail.nih.gov  
Oligo-dT track not found, Not 1 site shown in beginning of sequence  
is likely internal to the message. cDNA Library Preparation: M.B.  
Soares Lab Clone distribution: Researchers may obtain BMP4 cDNA  
clones from RESEARCH GENETICS. It should be noted that Bento Soares  
is generating a small number of additional specialized  
non-redundant arrays of BMP4 cDNAs whose availability will be  
considered under appropriate and limited collaborative arrangements.  
The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
Yale University School of Medicine  
Seq primer: M13 Forward  
POLYA-MO.



Oy	300	AlarfyatlglyleuglylethThrvalluPhurmettrrthngInSerSerclyser	312
Db	313	GCACGATTGCC-CTGGGCATCACACACAGCTCTGACATGACTACACAGTTCCAGTTCC	371
Oy	320	ArglaSerleupProLyValSerTyrrValLysAlaIleaspiIetrrpMetAlaValCys	339
Db	372	AGGCA-TCCTCGCCAAAGGCTCTCTATGTGAACCAATGTGACATCTGATGGCGGTATGC	430
Oy	340	LeuieupheValPheAlaAlaLeuIeugluIurrrAlaAlaIeasPhaValSerArgIn	359
Db	431	CTCTCTTTGGTTGGTGGCTCTACTGTGAAATATGACGAGTGAACCTT-GTCTCCAGGCAA	489
Oy	360	HisLysgluPheIleArgLeuArgArgArgInIurArgGlnArgLeuIuGluAspIle	379
Db	490	CATTAAGAGTTCCTCTCGTCCGAGACGACAGAAAGACGAGTAATTAAGAAAGATGTT	549
Oy	380	IleGlnIuSerArgPheTyrrPheArgGlyIrrGlyLeuGlyHisGlyLeuGlnAlaArg	399
Db	550	ACTGCTGAATACGTTTAACTTAACTGATGGCTATAGGGATGGGTACTGCTCCATGAATA	609
Oy	400	AspGlyGlyPromeTGLuGlySercGlyIleTyrrSer---ProGlnProPro	415
Db	610	GATGGCAAGCTGTCAAGGCTACACACTGCCCAACCACTTCACCAACCCCA	660
RESULT 9			
LOCUS	B0938794	975 bp	mRNA
DEFINITION	AGNCOCURT_8930535 NIH_MGC_94	Mus musculus	CDNA clone IMAGE:646531
ACCESSION	B0938794		
VERSION	B0938794.1	GI:22354272	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 975)		
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: <a href="mailto:cga@bbs-riemail.nih.gov">cga@bbs-riemail.nih.gov</a>		
	Tissue Procurement: The Cepko Laboratory		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	Clone Sequencing by: Agencourt Bioscience Corporation		
	CDNA distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>		
	Plate: LLM13990 row: p column: 20		
	High quality sequence stop: 658.		
FEATURES			
Source	Location/Qualifiers		
	1..975		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:646531"		
	/clone_lib="NIH MGC_94"		
	/tissue_type="retina"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: eye; Vector: pCMV-Sport6; Site:1: NotI;		
	Site:2: SalI; Cloned unidirectionally; oligo-dT primed.		
	Average insert size 3.3 kb. Library enriched for		
	full-length clones and constructed by Life Technologies.		
	Note: this is a NIH_MGC library."		
BASE COUNT	263 a 234 c 221 g 257 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	8,54e-75	length:	975
Score:	672.50	Matches:	144
Percent Similarity:	67.59%	Conservative:	52
Best Local Similarity:	49.66%	Mismatches:	67
Query Match:	29.86%	Indels:	27

[illegible]

SOURCE  
Mus musculus (strain:C57BL/6J) adult male hippocampus cDNA to mRNA.  
Clone:J1b0R1REN full-length enriched mouse cDNA library  
Clone:2900059A15.

ORGANISM  
Mus musculus

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

TITLE  
Carninci, P. and Hayashizaki, Y.  
High efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL  
MEDLINE  
PUBMED  
99279253  
10349636

REFERENCE  
AUTHORS  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL  
MEDLINE  
PUBMED  
20499374  
11042159

REFERENCE  
AUTHORS  
3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,  
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubiki, M.,  
Yoneeda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
R1EN Integrated sequence analysis (RISA) system -384 format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED  
20530913  
11076861

REFERENCE  
AUTHORS  
4  
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, O., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadoi, K., Matsuda, H., Ashburner, M., Batalov, S., Casavani, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiya, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L.M., Struhl, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hochmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
Toyo-Oka, K., Wang, K.H., Weltz, C., Whitaker, C., Wilming, L.,  
Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Koltschuk, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

JOURNAL  
MEDLINE  
PUBMED  
21085660  
11217851

REFERENCE  
AUTHORS  
5 (bases 1 to 2475)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hangaki, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F.,  
Hume, D., Imocanli, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kunihira, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.  
Direct Submission

JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGACGACAGACATCCAGACTCTTTTTCCTTTTTTTTTCN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAAGAGAGACTTCGATTAAATTAATTAATTCGCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR.

FEATURES

Source

Location/Qualifiers

1..2475

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="FANTOM,DB:2900059A15"

/db\_xref="MGD,MGI:1905909"

/db\_xref="taxon:10090"

/clone="2900059A15"

/sex="male"

/tissue\_type="hippocampus"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

1..2475

/gene="Gabra4"

501..2159

/gene="Gabra4"

/note="data source:MGI, source key:MGI:95616, evidence:ISS

gamma-aminobutyric acid (GABA-A) receptor, subunit alpha

4

putative"

/codon\_start=1

/protein\_id="BAB28975.1"

/db\_xref="GI:12851205"

/db\_xref="MGD:MGI:95616"

/translation="MVSQVQKPAALILSGVSLALLHFLCLAALESFGNSKDEKICPENPRRLDSLDGYNLRRCRGCPGVTEYTDITLVYSFGSVEMETRYNDVFROTWIDRLKTDGTIELRLNNMYTVPTPDTFRRKGSSVSHNMTPAKKLIRIMNGTTLTYTRLTIASCPMRWLDPMDGHACPKEGSVAVPSEMYVTWKGPENSVEVKPESLSLQVDLIGOVASSEIKSTGEYIVMTVFFHLRMKGFMIOITYIPCIINTVLISQVFSFWINKSEPVARTVGITTVLTJTTLISASHSLPKSYATAMDMFTAVCAFEVASLIIEFAVNVTETNIOMOKAKKKISKRPPEPPAPRVLYKEHTETSLONTHANLMRKRTNANLVHSEKSDSRTEVGNSSSKTSAOVESSEAPRAKHLASSPDRANAEEFMSAAAGGLSSASPSPHCTLPPASTLSGASSTRPAGCSRLGRKITTYNTTGGAAGNVSATPPPPAPPSSGSETSKIDKARLLPFVTFGAFNMVMVYVTLKSOTMEKSESLLM"

BASE COUNT

678 a 565 c 607 g 625 t

ORIGIN

Alignment Scores:

Pred. No.: 9 ABe-69 Length: 2475

Score: 631.50 Matches: 151

Percent Similarity: 51.29% Conservative: 68

Best Local Similarity: 35.36% Mismatches: 135

Query Match: 28.04% Indels: 73

Gaps: 11

DB:

US-10-075-846-4 (1-431) x AKO13727 (1-2475)

OY 7 AlathrLeuSerPhleuLeuLeuTrpThrLeuProGlyGlnValLeuLeuArgValAla 26  
111 111 111111 111  
555 GCCCTCGCATGCCTCCGTCCTG-----GGC 581





```

Db      117 AATGAAAAAGGCCAATTTTCATGATGACCCAGGAAAAACATCCTCTTATTTT 176
Oy      156 LysAsnGlyAsnValLeuTyrSerIleArgLeuThrIleLeuSerCysLeuMetAsp 175
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      177 CGGATGGAATGCTCTTGCAGCATGAGTTATCTATTACTCTTTATGCCCCTTTGAC 236
Oy      176 LeuLysAsnPheProMetAspIleGlnThrCysThrMetGlnLeuGlnSerSerIle 195
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      237 TTGACATGTTTCCCATGATACACACGTTGCAAGATGCACACGAG----- 284
Oy      196 LeuCysSerProLeuProSerLeuSerLeuSerValGlyTyrThrMetLysAspLeuVal 215
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      285 -----AGCTTTGTTACACACAACTGATTTACGA 314
Oy      216 PheGlnTyrPLeuGlnAspAlaProAlaValGlnValAlaGlnGlyLeuThrLeuProGln 235
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      315 TTTATCTGCAGTCAGCAGATCC--GTGCAATTA--GAAAAAATTCCTTCCCTCA 368
Oy      236 Phe---IleLeuArgAspGlyLysAspLeuGlyCysCysThrLysHisTyr---AsnThr 253
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      369 TTGATATCAAAAAGAAAGATATTGATATGTAACGTACAAATACTATAAGGACG 428
Oy      254 GlyLysPheThrCysIleGlnValLysPheHisLeuGlnMetGlyTyrTyrLeu 273
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      429 GACAAACAAAMADBKTKGTATCATCTCACCTGAGAGGAGGCGCTTTTACATG 488
Oy      274 IleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTyrValSerPheTyrIle 293
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      489 ATGGGGGTACAGCCCAACCCCTCATCTTCTCTCTCTGCTTCTTCTTGATC 548
Oy      294 AsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThrValLeuThrMetThr 313
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      549 AACCCGGACGAGTGTCTGCAGAGTCCCTGCGTATCTTCTCAGTCCTCAGCTTGCC 608
Oy      314 ThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyrValLysAlaIleAsp 333
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      609 TCTGATGCAACACCTTGCCTGAGCTTCCCAAACTTTCCTTGTGAAGCTCTGAT 668
Oy      334 IlePheMetAlaValCysLeuLeuPheValPheAlaLeuLeuGlnTyrAlaAlaIle 353
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      669 GTTGGCTTATGCTTCTCTCTTCTTGGTTTCTTCCCTGCGAGATGCGATTGTC 728
Oy      354 -----AsnPheValSerArgGlnHisLysGlnPheIleArgLeuArgArg 369
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      729 CAGGTGATGTCGACACCCCAAAAGGTTGAAGCTGAAAAAGCCAGATTGCTAAAGGCT 788
Oy      370 GlnArg 371
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      789 GAGCAA 794

```

```

RESULT 12
LOCUS   B0042344      802 bp      mRNA      linear      EST 28-MAR-2002
DEFINITION
  UT-M-EH0-bug-p-09-0-UI.r2 NIH-BMAP-EH0 Mus musculus cDNA clone
  IMAGE:5685752 5', mRNA sequence.
ACCESSION
  B0042344
  B0042344.1 GI:19791298
VERSION  EST.
KEYWORDS
  house mouse.
SOURCE   Mus musculus.
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    1 (bases 1 to 802)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaabs@email.nih.gov
  Tissue Procurement: Dr. James Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

Location/Qualifiers

```

1. 802
   /organism="Mus musculus"
   /strain="C57BL/6"
   /db_xref="taxon:10090"
   /clone="IMAGE:5685752"
   /clone_1b="NIH-BMAP-EH0"
   /tissue_type="whole brain"
   /dev_stage="embryo 18.5 dpc"
   /lab_host="DH10B (TI phage resistant)"
   /note="Organ: brain; Vector: PYX-Asc; Site_1: EcoR I;
   Site_2: Not I; The library was constructed according to
   Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
   1996. Denatured mRNA was size fractionated on a 1% agarose
   gel. First strand cDNA synthesis was primed with an
   oligo-dT primer containing a Not I site. Double stranded
   cDNA was size selected according to mRNA size fraction,
   ligated with EcoR I adaptor, digested with Not I, and then
   cloned directionally into PYX-Asc vector. The library tag
   sequence located between the Not I site and the polyA tail
   is CAGCCAGCAGC. This library was created for the
   University of Iowa Mouse Brain Molecular Anatomy Project
   (BMAP). Gene discovery in the developing Mouse Nervous
   System, supported by National Institute of Mental Health
   (NIMH), Hemlin Chin, Ph.D., program coordinator."

```

BASE COUNT 237 a 167 c 174 g 224 t

ORIGIN

Alignment Scores:

```

Pred. NO.: 3.79e-68      Length: 802
Score: 619.50      Matches: 136
Percent Similarity: 62.32%      Conservative: 41
Best Local Similarity: 47.89%      Mismatches: 62
Query Match: 27.51%      Indels: 45
DB: 14      Gaps: 9

```

US-10-075-846-4 (1-431) x B0042344 (1-802)

```

Oy      21 ValLeuLeuArgValAlaLeuAlaLysGlnGluValLysSerGlyThrLysGlySerGln 40
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      5 TTGTTGTTTGAAGACCGCTTGCTGTAAGAAAAGTCTTCCAAGAAAGGAAAGGAAAAAG 64
Oy      41 -----Promet 42
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      65 AAGCAGTACTTGTGCCCATCTCACAGCTCACCCGACGACCTGGCCGTTGCCCCCAAC 124
Oy      43 SerProSerAspPheLeuAspLysLeuMetClyArgThrSerGlyTyrAspAlaArgIle 62
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      125 TTCACGACCAATATCTTGAACAGGCTGCTG-----GTCACTTATGATCCCAAGATC 175
Oy      63 ArgProAsnPheLysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPhe 82
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      176 AGACCAACTTCAAGAGCATTCTGTTGATGATAGTACCAACATTTTATTATAAGTTT 235
Oy      83 SerSerValThrLysThrThrMetLysPtyrArgValAsnValPheLeuArgGlnGlnTyr 102
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      236 GATCCATTTCAGAGACACAACTGACTATTAACATTTCTTGACAGACAGAAATGG 295
Oy      103 AsnAspProArgLeu-----SerTyrArgGlyTyrProAspAspSerLeuAsp 118
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      296 AATGACCCACAGCTCAAGTACTAGTACTCAGA-----GGCTCAGATGCACCTACA 349
Oy      119 LeuAspProSerMetLeuAspSerIleTyrPtyrProAspLeuPhePheAlaAsnGlyLys 138
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      350 GTTACCCCAACCATGTATAGTCTTGTGGAACCTGACTTATTTTGCAAAATGAAAA 409
Oy      139 GlyAlaAsnPheHisGlnValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGly 158
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db      410 AGTGCCATTTCATGATGTCGACCCAGAAATATCTGTGTATCTTTCGGATGGA 469
QY      159 AsnValleuTySerIleArgleuThrIleuSerCysLeuMetAspIleuysAsn 178
        ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      470 GACGCTCTTGAGGATGAGGTGTCTATACACTTCACTGCTCCTGACTTAACCTGTG 529
QY      179 PheProMetAspIleuThrCysThrMetGlnleuGluSerSerIleuGlySer 198
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      530 TTTCCCATGACACACACACGCTGCAAAATGCAACTTGAG----- 568
QY      199 ProleuProSerLeuSerLeuSerValGlyTyThrMetLysAspLeuValPheGluTrp 218
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      569 -----AGCTTGTGATATACAAACCGAGATTAGATTACATCTCAG 607
QY      219 LeuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPhe---Ile 237
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      608 CAGTACAGAGATCCCT---GTTCAAGTTG---GAAAAAATTGCTTACCTCAATTATCATATT 661
QY      238 LeuArgAspGluLysAspLeuGlyCysCysThrLysHisIleTyrr---AsnThrGlyLysPhe 256
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      662 AAAAAGAGAGATATGCAATATGCGCAACTGCAAAATCTATAAGGACACTGCTACTAC 721
QY      257 ThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyrrTyrrLeuIleGlnMet 276
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      722 ACTTGTGTGGAGGTCTCTTACCTGAGAGACAGGTGGTGTCTACATGATGGCGCTA 781
QY      277 TyrrIleProSer 280
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      782 TATGCACCACT 793

RESULT 13
AUI33223      817 bp  mRNA  linear  EST 01-AUG-2002
LOCUS      AUI33223 NT2RP4 Homo sapiens cDNA clone NT2RP401569 5', mRNA
DEFINITION      sequence.
ACCESSION      AUI33223
VERSION      AUI33223.1 GI:10993762
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 817)
                Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y., Saito,K., Yamamoto,J.,
                Nishikawa,T., Nakamura,Y., Nagai,T., Sugano,S., Masuno,Y. and
                Isogai,T.
                HRI human cDNA project (Ota,T., Sugiyama,T., Ishii,S., Suzuki,Y.,
                Saito,K., Yamamoto,J., Nishikawa,T., Nakamura,Y., Nagai,T., Sugano
                S., Masuno,Y., Isogai,T.)
                Unpublished (2000)
JOURNAL      Contact: Takeo Isogai
                Genomics Laboratory
                Helix Research Institute
                1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
                Tel: 81-438-52-3875
                Fax: 81-438-52-3986
                Email: genomicehri.co.jp
                HRI human cDNA project; 5', & 3'-end one pass sequencing; Helix
                Research Institute; cDNA library construction; Department of
                Vitrology, Institute of Medical Science, University of Tokyo, and
                Helix Research Institute.
FEATURES
    source
        1..817
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="NT2RP4001569"
            /clone_lib="NT2RP4"
            /cell_type="teratocarcinoma"
            /cell_line="NT2"
            /note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
            cells after 2-weeks retinoic acid (RA) induction"
BASE COUNT      207 a 207 c 198 g 202 t 3 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      4,866-67
Score:          611.00
Percent Similarity: 62.14%
Best Local Similarity: 43.21%
Query Match:    27.13%
DB:
      9
      Gaps: 3

US-10-075-846-4 (1-431) x AUI33223 (1-817)
QY      57 GlyTyrrAspAlaIleArgProAsnAspHisGlyProProValAsnValThrCysAsn 76
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      15 GCGTACGACATTCGCGCTTAAGACCCGACTTCGGGGGTCCCGGCTCGGTGGGATGAC 74
QY      77 IlePheIleAsnSerPheSerSerValThrLysThrThrMetAspTyrrAlaValAsnVal 96
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      75 ATGCACATCGCCGACATCGACATGCTTCCGAGTCAACATGATGATTAACTTAACCATG 134
QY      97 PheLeuArgGlnIleTrpAsnAspProAlaArgLeuSerTyrrArgGluTyrrProAspAspSer 116
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      135 TATTTCACACATATTGAGAGATTAAGAGCTCGCTATTTCTGGATCCCT---CTCAC 191
QY      117 LeuAspLeuAspProSerMetLeuAspSerIleTrpLysProAspLeuPheAlaAsn 136
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      192 CTCACGCTTGACAAATCGAGCTGGCTGACCAAGCATGAGGCGCCGACACATATTCTTAAT 251
QY      137 GluLysGlyAlaAsnPheHisGlnValThrThrAspAsnLysLeuLeuArgIlePheLys 156
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      252 GACAAAGATCATTTGTGCTAGTGAAGTGAACCTGAAACCCGATGATCCGCTTCACCT 311
QY      157 AsnGluAsnValLeuTyrrSerIleArgleuThrIleuSerCysLeuMetAspLeu 176
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      312 GATGGAGACAGTCTGTATGGCTCGAGATTCACACAGACAGCATGATGATGACCTC 371
QY      177 LysAsnPheProMetAspIleGlnThrCysThrMetGlnleuGluSerSerIleLeu 196
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      372 AGGAGATACCCCTGAGACGACAGACACGACTCGAATATTGAA----- 416
QY      197 CysSerProleuProSerLeuSerLeuSerValGlyTyrrTrpMetLysAspLeuValPhe 216
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      417 -----AGCTATGGCTACACACAGCATGATGATGAGATT 449
QY      217 GluTyrrLeuGluAspAlaProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPhe 236
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      450 TACTGCGAGGCGGGGACAGACGCTGTACCGAGTGAAGAGATTGACTCCCGAGTTC 509
QY      237 IleLeuArgAspGluLysAspLeuGlyCysCysThrLysHisIleTyrrAsnThrGlyLysPhe 256
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      510 TCCATCGTGGAGACACCT---CTGCTTCGAGGATGTTGCTTCGCCACAGGTCCTAT 566
QY      257 ThrCysIleGluValLysPheHisLeuGluArgGlnMetGlyTyrrTyrrLeuIleGlnMet 276
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      567 CCGTCACTGTACGAGCTTTCGGTTGAGAGGAAACATTGATCTTCAATCTTCAGACT 626
QY      277 TyrrIleProSerLeuLeuIleValIleuSerTrpValSerPheTrpIleAsnMetAsp 296
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      627 TATATGCCCTCATATCTATACAGATTTCTGTCGGGGGTCTCTTGATCAATTATGAT 686
QY      297 AlaAlaProAlaArgValGlyLeuGlyLysThrThrValLeuThrMetThrThGlnSer 316
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      687 GCATCTGCTGTAGAGTTGCTCGCGGATCAACATGCGGCTGACAAATGACCAACATTNAC 746
QY      317 SerGlySerArgAlaSerLeuProValValSerTyrrValIleAlaIleAsnIleTrpMet 336
        ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      747 ACCCACTTCGAGACACCTTGCACCAATNCCCTATGTCAAAGCATTTGACATGATCTT 806

RESULT 14
BC029850      1942 bp  mRNA  linear  HTC 20-MAY-2002
LOCUS      BC029850
DEFINITION      Homo sapiens, gamma-aminobutylic acid (GABA) A receptor, gamma 3,
                clone IMAGE:5178575, mRNA.
ACCESSION      BC029850

```

VERSION BC029850.1 GI:20988273  
 KEYWORDS HTC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1942)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-MAY-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgapbs-ref@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNI)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amge@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LILNI at: http://image.lilni.gov  
 Series: IRAC Plate: 51 Row: 0 Column: 23  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 15193297  
 This clone has the following problem: frame shifted.

## FEATURES

## SOURCE

1..1942  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="locusID:2567"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5178575"  
 /tissue\_type="Brain, Lung, Testis, adult, pooled whole"  
 /clone\_lib="NIH\_MGC\_115"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

BASE COUNT 545 a 492 c 435 g 470 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 1.51e-65 Length: 1942  
 Score: 605.00 Matches: 143  
 Percent Similarity: 54.64% Conservative: 63  
 Best Local Similarity: 37.93% Mismatches: 129  
 Query Match: 26.87% Indels: 43  
 DB: 11 Gaps: 7

US-10-075-846-4 (1-431) x BC029850 (1-1942)

OY 6 ProAaThrLeuSerPheLeuLeuLeuTrpThrLeuProGlyGlnValLeuLeuArgVal 25  
 ||| ||||| ||| ||||| :  
 DB 134 CCGCGGACCTCGCGCCGAGCTCCACGACACATGCGCCGGAAGCTCTGCTCTCCTC 193  
 OY 26 AlaLeuAlaLys-----GlnGlu 31  
 ||| ||||| ||| ||||| :  
 DB 194 TGCGTGTCTCGGGCTTCACGCGCGGTCCAGAAAGTGAAGAGATGAATATGAGAT 253  
 OY 32 ValLysSerGlyThrLys-----GlySerGlnProMetSerProSerasp 46  
 ||| ||||| ||| ||||| :  
 DB 254 TCATCATCAACCAAGTGGGTCTTGCTCCAAATCTCCAAAGACACCGACGCTGCTCTT 313  
 OY 47 PheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgIleArgProAsnhe 66  
 ||| ||||| ||||| ||||| :  
 DB 314 ATTCTCAACAAAGTCTCTAAGACA-----TATGATAAAAAGCTGAGCGCATATATT 364

OY 67 LysGlyProProValAsnValThrCysAsnIlePheIleAsnSerPheSerSerValThr 86  
 ||| ||| :  
 DB 365 GGAATTAACCGACCGCATATTGACGCTTACATTTATGTAACAGATGTCGCTGTGCA 424  
 OY 87 LysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnIleTrpAsnProArg 106  
 ||| ||||| ||||| ||||| :  
 DB 425 TCAATTAACATGGAATGCAATGCAATGCAATGCTTGGT-CAGACCTGAGAGATAGTCGC 483  
 OY 107 LeuSerTyrArgGluTyrProAspAspSerIleuAspLeuAspProSerMetLeuAspSer 126  
 ||| ||| :  
 DB 484 CTTCGATTCACACAGC---ACAATGAATAATCTTACTGCAACAGCAGCATGTCGGGCTTA 540  
 OY 127 IleTrpLysProAspLeuPheAlaAsnGlnLysGlnGlyAlaAsnPheIleGlnValThr 146  
 ||| ||||| ||||| ||| ||| :  
 DB 541 ATCTGATCCAGACACACATCTCCGCAATTCATAAAGGAGAGGCTCATGTGATGAC 600  
 OY 147 ThrAspAsnLysLeuLeuArgIlePheLysAsnGlnLysAsnValLeuTyrSerIleArgLeu 166  
 ||| ||||| ||||| ||||| :  
 DB 601 ACACCAATCAGCTCTCCGATTTGGAAATGACGGAATAATCTTTACACTTTCAGGCTC 660  
 OY 167 ThrIleuIleLeuSerCysLeuMetLysAsnLysAsnProMetAspIleGlnThrCys 186  
 ||| ||| :  
 DB 661 ACCATCAATGCTGAGTGCACGCTGACGCTGCACACATTCCTCCATGAGCAACACTCTGC 720  
 OY 187 ThrMetGlnLeuGlnSerSerIleLeuCysSerProLeuProSerLeuSerLeuSer 206  
 ||| ||| :  
 DB 721 -----CCGCTGATTTTCTCCAGC 738  
 OY 207 ValGlyTyrThrMetLysAspLeuValPheGlnTrpLeuGlnAspAlaProAlaValGln 226  
 ||||| ||||| :  
 DB 739 TATGCTATCCCAAGAAAGAAATGATTAATGATGCAAAAAAATTCAGTGAGACACT 798  
 OY 227 ValAlaGlnGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLysAspLeuGlnCys 246  
 ||| ||||| :  
 DB 799 GACCAGAAATGATGCGGCTTATTCATGATT-----GACTTCATGCGCTCAGAAAC 849  
 OY 247 CysThrLysHisTyrAsnThr-----GlyLysPheThrCysIleGlnValLysPheHis 264  
 ||||| ||| ||| :  
 DB 850 ACCACAGAAATCGTACACACGTCGACGATGATTATGTCATGACATCATATATTGTA 909  
 OY 265 LeuGlnArgGlnMetGlyTyrTyrLeuIleGlnMetCysTrpIleProSerLeuLeuIleVal 284  
 ||| ||||| ||||| ||||| :  
 DB 910 TTGAGTAAAGAAATGGATGACTTCCACATGACATGATCATCTCTGATCTGCTGTG 969  
 OY 285 IleLeuSerTrpValSerPheTrpIleAsnMetAspAlaIleProAlaArgValGlyLeu 304  
 ||||| ||||| ||||| ||||| :  
 DB 970 GTTTATCTCGGTGCTCATTTGGATCAAAAAAGATGCTACGCCAGAAACACATTA 1029  
 OY 305 GlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuPro 324  
 ||||| ||||| ||||| ||||| :  
 DB 1030 GGCATCACACGAGTGTGACCATGACCAACCTGAGACCATCGCAGGAGACGCTTGCCA 1089  
 OY 325 LysValSerTyrValLysAlaIleAspIleThrMetAlaValCysLeuLeuPheValPhe 344  
 ||||| ||||| ||||| ||||| :  
 DB 1090 CGCGTGTCTACGTACCGCATGACCTTTTGTGATGCTGTCTCTCTGCTCTCTC 1149  
 OY 345 AlaAlaLeuLeuGlnTyrAlaAlaIleAsnPheValSerArgGlnHisLys 361  
 ||||| ||||| ||||| ||||| :  
 DB 1150 GCCGCGTGAATGAGATATGCAACCTCAACTACTTATTCACAGCTGAGAAA 1200  
 RESULT 15  
 AL539373 948 bp mRNA linear EST 16-FEB-2001  
 LOCUS AL539373 LTI.FL013.FBm1 Homo sapiens cDNA clone CS0DF033YML7 5  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL539373  
 VERSION AL539373.1 GI:12868529  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 948)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
 source 1. 948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CSDF033YM17"  
 /clone\_lib="LTI\_F1013\_FBrn1"  
 /dev\_stage="pooled tissue from post conception fetuses (20  
 week, 24 week and 26 week)"  
 /lab\_host="DH10B"  
 /note="Organ: Fetal brain; Vector: pCMVSPORT 6; 1st strand  
 cDNA was primed with a NotI-oligo(dt) primer. Five prime  
 end enriched, double-stranded cDNA was digested with Not I  
 and cloned into the Not I and Eco RV sites of the  
 pCMVSPORT 6 vector. Library was constructed by life  
 technologies. Contact : Feng Liang Life Technologies, a  
 division of Invitrogen 9800 Medical Center Drive Rockville  
 , Maryland 20850, USA Fax : (1) 301 610 8371 Email :  
 fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"

BASE COUNT 275 a 215 c 210 g 247 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.03e-61 Length: 948  
 Score: 570.50 Matches: 121  
 Percent Similarity: 57.64% Conservative: 60  
 Best Local Similarity: 38.54% Mismatches: 108  
 Query Match: 25.33% Indels: 25  
 DB: 9 Gaps: 6

US-10-075-846-4 (1-431) x AL539373 (1-948)

```

Oy 10 SerPheLeuLeuLeuThrPheProGlyGlnVal-----LeuLeuArgValAla 26
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 63 TCAGAGTTAGTATGTCGACAGTACAAATCGAGAGACTGCTGGGCTTCTCTTCCT 122
Oy 27 LeuAlaLysGluGluValLysSerGlyThrLysGlySerGlnProMetSerProser--- 45
    ::| ::|:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 123 GTGATGATTACCATGCTGTGTGTCACACACACACCAATGAACCCAGCAATGTCATAC 182
Oy 46 -----AspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgIleArg 63
    ::| ::|:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 183 GTGAAGAGACAGTGTGACGATTGCTC-----AAAGGATATGACATTGCTTGGCGG 233
Oy 64 ProAsnPhelYsGlyProProValAsnValThrcysAsnIlePheIleAsnSerPheSer 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 234 CCGGACTTGGAGGAGGCCCCCTCGACGCTGGGATGCGGATGATGTCGCCAGCATAGAC 293
Oy 84 SerValThrLysThrThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsn 103
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 294 ATGGTCTCCGAGAGTGAATGATATACATCACCATGATATTTCCAGCAGCTTGGAAGA 353
Oy 104 AspProAlaGluSerTyrArgIleuTyrProAspAspSerLeuAspLeuAspProSerMet 123
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||
Db 354 GACAAAGAGCTTCTTATTCTGAAATCCCA--CTGAACCTCACCTTACACAAATAGGCTA 410
Oy 124 LeuAspSerIleTrpLysProAspLeuPhePheAlaAsnGluLysGlyValAsnPheHis 143
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||
Db 411 GCTGACCAACTCTGGTACCAGACACTTCTTGAATGACAGAAATCATTTGTGCAT 470
Oy 144 GluValThrThrAspAsnLysLeuLeuArgIlePheLysAsnGlyValAsnValLeuTyrSer 163
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||
Db 471 GGGGTACACAGTGAATAATGATGATGACATCCATGATGGAACAGTTCTCTATGGA 530
Oy 164 IleArgLeuThrLeuLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIle 183
    ::| |||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

Db 531 CTCGCAATCACAACCAACAGCTCATGATGATGATCTTCGAAGATATCCATGATGATGAG 590
Oy 184 GlnThrCysThrMetGlnLeuGlnUserSerSerIleLeuCysSerProLeuProSerLeu 203
    ||| |||:|||||:|||||:|||||:|||||:|||||:|||||
Db 591 CAGAACTGCACCCCTGAGATCGAA----- 614
Oy 204 SerLeuSerValGlyTyrThrMetLysAspLeuValPheGluTrpLeuGlnAspAlaPro 223
    ||| |||:|||||:|||||:|||||:|||||:|||||
Db 615 -----AGTTATGGCTATACCACTGATGACATTTTACTGCAATGAGAGAGAGAGG 668
Oy 224 AlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGlnLysAsp 243
    ||| |||:|||||:|||||:|||||:|||||:|||||
Db 669 GCAGTCACACTGCTGTAATTAATTCGAACACTCTCTCAATTTTCAATTGTGACAAAGATG 728
Oy 244 LeuGlyCysCysThrLysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPhe 263
    ::| ::|:|||||:|||||:|||||:|||||:|||||
Db 729 GTGCTTAAGAGAGTGGAG--TTCAACACAGAGGCGTATCCAGCACTGCATTAAGTTT 785
Oy 264 HisLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIle 283
    |||:|||||:|||||:|||||:|||||:|||||:|||||
Db 786 CGCTTAAGAGAGAAACATTGCTTACTCATTTTGCAAACCTACATGCTTCTACACTGATT 845
Oy 284 ValIleLeuSerTyrValSerPheTrpIleAsnMetAspAlaAlaProAlaArgValGly 303
    ||| |||:|||||:|||||:|||||:|||||:|||||
Db 846 ACAATTCTGTCTCTGGGTGCTTTTGGATCAACATGATGATCTGCAGACCGAGCTCCCA 905
Oy 304 LeuGlyIleThrThrValLeuThrMetThrThrGlnSerSer 317
    ||| |||:|||||:|||||:|||||:|||||
Db 906 CTAGGATTCGACAKATCTTTTACAATGACAAACATCAGCAC 947

```

Search completed: July 1, 2003, 00:55:27  
 Job time : 2277.88 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 22:23:50 ; Search time 131.866 Seconds  
(without alignments)  
3531.204 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75

Sequence: 1 SSSILCSPLPSLSLV 16

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPRO\_spool/US10075846/runat\_25062003.163648.5147/app\_query.fasta\_1.782  
-DB=genEmbl -QFRT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -List=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846.ecgn.1.1.3854.qrunat.25062003.163648.5147 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEBUFFER -NEG.SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vl:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_dln:\*  
35: em.htg\_fod:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	90.7	39796	6 AX037570	AX037570 Sequence
2	68	90.7	39796	9 HSU35C3	Z93848 Human DNA s
3	61	81.3	170676	9 AC026258	AC026258 Homo sapi
4	61	81.3	183125	9 AP002848	AP002848 Homo sapi
5	61	81.3	189127	2 AC084114	AC084114 Homo sapi
6	61	81.3	195477	2 AC079196	AC079196 Homo sapi
7	60	80.0	91582	2 AC109887	AC109887 Rattus no
8	57	76.0	169845	2 AC095732	AC095732 Rattus no
9	57	76.0	191204	2 AC094926	AC094926 Rattus no
10	56	74.7	104525	2 AC095220	AC095220 Rattus no
11	56	74.7	172555	2 AC102866	AC102866 Mus muscu
12	55	72.3	598	5 G6126MRP	X61200 G. gallus 12
13	55	72.3	139754	2 AC118327	AC118327 Rattus no
14	55	72.3	164772	2 AC095184	AC095184 Rattus no
15	55	72.3	174354	2 AC108830	AC108830 Mus muscu
16	55	73.3	193627	2 AC091427	AC091427 Mus muscu
17	55	73.3	197679	2 AC114635	AC114635 Mus muscu
18	55	73.3	217164	2 AC094182	AC094182 Rattus no
19	55	73.3	238033	2 AC125309	AC125309 Mus muscu
20	55	73.3	254752	2 AC127336	AC127336 Mus muscu
21	55	73.3	310148	2 AC073729	AC073729 Mus muscu
22	54	72.0	309	9 HS255ZEP	Z17151 H. sapiens
23	54	72.0	14839	1 AE000738	AE000738 Aquifex a
24	54	72.0	69740	2 AC087748	AC087748 Homo sapi
25	54	72.0	85149	9 AL137879	AL137879 Human DNA
26	54	72.0	106186	2 AC105876	AC105876 Rattus no
27	54	72.0	113704	9 HS267P19	Z75889 Human DNA s
28	54	72.0	117628	9 AC008891	AC008891 Homo sapi
29	54	72.0	120652	2 AL353724	AL353724 Human DNA
30	54	72.0	121705	2 AC118360	AC118360 Rattus no
31	54	72.0	153042	9 AP002984	AP002984 Homo sapi
32	54	72.0	153788	2 AC118805	AC118805 Rattus no
33	54	72.0	164472	9 AC018897	AC018897 Homo sapi
34	54	72.0	178273	2 AC068786	AC068786 Homo sapi
35	54	72.0	179217	2 AC099748	AC099748 Bos tauru
36	54	72.0	180226	2 AC109001	AC109001 Rattus no
37	54	72.0	209621	2 AC111125	AC111125 Mus muscu
38	54	72.0	223150	2 AC104396	AC104396 Mus muscu
39	54	72.0	233454	2 AC096987	AC096987 Rattus no
40	54	72.0	263744	2 AF401201	AF401201 Mus muscu
41	53.5	71.3	35412	8 SPAC4F8	Z98530 S. pombe chr
42	53.5	71.3	174005	2 AC117603	AC117603 Mus muscu
43	53	70.7	549	4 AB048310	AB048310 Equus cab
44	53	70.7	3301	6 AX333072	AX333072 Sequence
45	53	70.7	3301	9 H0MPPARA	L07592 Human perox

#### ALIGNMENTS

RESULT 1

AX037570  
LOCUS AX037570 39796 bp DNA linear PAT 16-NOV-2000  
DEFINITION Sequence 6 from Patent WO0058461.  
ACCESSION AX037570  
VERSION AX037570.1 GI:11226991  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Rappold-Hoerbrand, G.  
TITLE Gene for ataxia  
JOURNAL Patent: WO 0058461-A 6 05-OCT-2000;  
RAPPOLD HOERBRAND GUERIN (DE)  
FEATURES  
source  
1. 39796  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 10630 a 9337 c 9355 g 10474 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.92 Length: 39796  
Score: 68.00 Matches: 14  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 93.33% Mismatches: 0  
Query Match: 90.67% Indels: 0  
DB: 6 Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AX037570 (1-39796)  
QY 2 SerSerIleuGysSerProLeuProSerLeuSerLeuSerVal 16  
Db 35719 GCATCATACTCTGCGAGCCCTCGCATCTCTCTCACTTCACTT 35763  
RESULT 2  
HS03563/c 39796 bp DNA linear PRI 23-NOV-1999  
LOCUS HS03563/c  
DEFINITION Human DNA sequence from cosmid U3563, between markers DXS366 and  
ACCESSION DXS87 on chromosome X contains ESTs.  
VERSION 293848  
KEYWORDS 293848.1 GI:1934909  
X.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Burgess, J.  
TITLE Direct Submission  
JOURNAL Submitted (07-APR-1997) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1RO, UK. E-mail enquiries: humquery@sanger.ac.uk  
COMMENT  
IMPORTANT: This sequence is the entire insert of clone U3563. This  
sequence has been finished according to sequence map criteria as  
follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known  
annotated human repeat sequence elements (e.g. Alu). Where the  
sequence is ambiguous, there is an annotation using the 'unsure'  
feature key.  
The true left end of clone U3563 is at 1 in this sequence. The true  
right end of clone U3563 is at 39796.  
U3563 is from the Lawrence Livermore National Laboratory  
flow-sorted X chromosome cosmid library LLOXNC01.  
FEATURES  
source  
1. 39796  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="X"  
/clone="LLOXNC01-3563"  
/clone\_11b="LLOXNC01"

repeat\_region  
9. .107  
/note="MIR2 repeat: matches 48. .146 of consensus"  
repeat\_region  
1267. .1477  
/note="MIR repeat: matches 1. .218 of consensus"  
repeat\_region  
2267. .2446  
/note="MIR repeat: matches 196. .19 of consensus"  
repeat\_region  
2619. .2734  
/note="L1P6 repeat: matches 571. .457 of consensus"  
repeat\_region  
2863. .3051  
/note="MIR5A repeat: matches 7. .189 of consensus"  
repeat\_region  
3608. .3745  
/note="MIR repeat: matches 20. .156 of consensus"  
repeat\_region  
6949. .7115  
/note="MIR repeat: matches 212. .46 of consensus"  
repeat\_region  
7944. .8166  
/note="MIR repeat: matches 224. .2 of consensus"  
repeat\_region  
8750. .8823  
/note="MIR repeat: matches 65. .138 of consensus"  
repeat\_region  
14745. .14874  
/note="MIR2 repeat: matches 146. .21 of consensus"  
repeat\_region  
15339. .15398  
/note="5 copies of 12 mer 82 % conserved"  
repeat\_region  
16847. .17148  
/note="AluSg repeat: matches 1. .302 of consensus"  
repeat\_region  
17947. .18235  
/note="AluYb repeat: matches 295. .1 of consensus"  
repeat\_region  
20312. .20454  
/note="MIR repeat: matches 20. .162 of consensus"  
repeat\_region  
20456. .20696  
/note="AluX repeat: matches 44. .284 of consensus;  
Incomplete repeat"  
repeat\_region  
21916. .22233  
/note="AluSp repeat: matches 1. .302 of consensus"  
repeat\_region  
22428. .22729  
/note="AluX repeat: matches 1. .302 of consensus"  
repeat\_region  
22816. .22915  
/note="MIR5A repeat: matches 71. .169 of consensus"  
repeat\_region  
23170. .23463  
/note="AluX repeat: matches 300. .2 of consensus"  
repeat\_region  
23578. .23716  
/note="MIR2 repeat: matches 16. .146 of consensus"  
repeat\_region  
23844. .23993  
/note="MIR1A1 repeat: matches 365. .215 of consensus"  
repeat\_region  
24402. .24703  
/note="AluYb repeat: matches 302. .11 of consensus"  
repeat\_region  
24710. .24794  
/note="MIR repeat: matches 110. .20 of consensus"  
repeat\_region  
25054. .25347  
/note="AluSg repeat: matches 299. .1 of consensus"  
repeat\_region  
25430. .25730  
/note="AluSp repeat: matches 303. .1 of consensus"  
repeat\_region  
25733. .26033  
/note="AluX repeat: matches 302. .2 of consensus"  
repeat\_region  
26644. .26753  
/note="FLAM\_A repeat: matches 123. .10 of consensus"  
repeat\_region  
26869. .27014  
/note="MIR repeat: matches 49. .196 of consensus"  
repeat\_region  
27270. .27486  
/note="MIR repeat: matches 31. .250 of consensus"  
repeat\_region  
27575. .27798  
/note="AluYb repeat: matches 87. .302 of consensus;  
Incomplete repeat"  
repeat\_region  
30559. .30679  
/note="AluSg repeat: matches 1. .121 of consensus;  
Incomplete repeat"  
repeat\_region  
30679. .30799  
/note="AluSg repeat: matches 124. .1 of consensus;  
Incomplete repeat"  
repeat\_region  
30706. .30781  
/note="L1MB8 repeat: matches 413. .676 of consensus"  
repeat\_region  
31110. .31403  
/note="AluSg repeat: matches 1. .293 of consensus"

```

repeat_region 31775..32116
/note="MIR1A1 repeat: matches 365. .1 of consensus"
repeat_region 32540..32840
/note="AluSp repeat: matches 1. .302 of consensus"
repeat_region 34108..34350
/note="MIR repeat: matches 247. .6 of consensus"
repeat_region 34410..34709
/note="AluSp repeat: matches 301. .1 of consensus"
repeat_region 34732..34831
/note="MIR repeat: matches 106. .203 of consensus"
repeat_region 34890..34976
/note="MIR repeat: matches 36. .122 of consensus"
repeat_region 36739..36891
/note="AluSp repeat: matches 148. .300 of consensus"
repeat_region 36958..37110
/note="MIR repeat: matches 105. .260 of consensus"
repeat_region 37468..38250
/note="MER21B repeat: matches 1. .770 of consensus"
prim_transcript <38771..>39796
/note="match: 5' EST C16803 clone GEN-530B01: match: 5'
EST C16720 clone GEN-522A09: match: 5' EST C1570 clone
GEN-549E01"
prim_transcript <38771..>39418
/note="match: 5' EST T86096 clone 114200: match: 3' EST
W67655 clone 343194"
BASE COUNT 10474 a 9355 c 9337 g 10630 t
ORIGIN
Alignment Scores:
Pred. No.: 9.92 Length: 39796
Score: 68.00 Matches: 14
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 0
Query Match: 90.67% Indels: 0
DB: 9 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x HS05G3 (1-39796)
Oy 2 SerSerIleuCySerProleuProSerIleuSerIleuSerVal 16
Db 4078 GCATTCATCTCTGTCAGCCCTCTGTCATCTCTGTCATCTTTCAGTT 4034
RESULT 3
AC026258 170676 bp DNA linear PRI 18-AUG-2001
LOCUS Homo sapiens chromosome 8, clone RP11-67N21, complete sequence.
DEFINITION AC026258
AC026258.4 GI:14579802
VERSION AC026258.4 GI:14579802
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 170676)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Unpublished
JOURNAL 2 (bases 1 to 170676)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choquet,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kam,L., Karatas,A.,
Klein,J., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,
Meldrim,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

```

```

O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessie,S., Theodore,J., Tirrell,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (21-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 170676)
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choquet,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gort,S.,
Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Menes,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nordu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Reta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Roselli,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnaz,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tessie,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
JOURNAL Submitted (18-AUG-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 1, 2001 this sequence version replaced gi:14336597.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7482
Center clone name: 67_N21
-----
FEATURES
SOURCE
Location/Qualifiers
1..170676
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-67N21"
/clone_lib="RPCI-11 Human Male BAC"
79..293
/rpt_family="MER58"
complement(913..1811)
/rpt_family="LIME2"
1929..2057
/rpt_family="LIME2"
2238..2911
/rpt_family="LIME2"
2916..2965
/rpt_family="AT-rich"
3230..3303
/rpt_family="L2"
/rpt_family="L2"
complement(3538..3699)
/rpt_family="MLTJ1"
complement(3924..4315)
/rpt_family="MSTB"

```

```
repeat_region      /rpt_family="MIR"      .27437)
                    /rpt_family="L2"
repeat_region      complement(.27520, .27655)
                    /rpt_family="MIR"
repeat_region      27834, .28144
                    /rpt_family="AluSc"
                    28499, 28616
repeat_region      /rpt_family="L2"
                    complement(.28831, .29045)
                    /rpt_family="MER20"
repeat_region      complement(.29201, .29416)
                    /rpt_family="L2"
repeat_region      30415, .30513
                    /rpt_family="GA-rich"
repeat_region      complement(.30784, .30992)
                    /rpt_family="MIR3"
repeat_region      complement(.31406, .31478)
                    /rpt_family="MER16"
repeat_region      33152, .33240
                    /rpt_family="L2"
repeat_region      complement(.33241, .33407)
                    /rpt_family="MER5A"
repeat_region      33408, .33810
                    /rpt_family="L2"
repeat_region      complement(.33832, .34249)
```

```

Alignment Scores:
Pred. No.:      506      length:      170676
Score:          61.00      Matches:      12
Percent Similarity: 93.33%      Conservative: 2
Best Local Similarity: 80.00%      Mismatches: 1
Query Match:      81.33%      Indels:      0
DB:              9      Gaps:      0

US-10-075-846-4_COPY_192_207 (1-16) x AC026258 (1-170676)
QY      1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15
      ::::::::::::::::::::::::::::::::::::::::::::
db      46678 ACACAATTCATTCGTGTCACCTCCACCCCAACCTAGGCCCTCT 46722

```

RESULT 4	
AP002848/c	
LOCUS	183125 bp DNA- linear PRI 20-JAN-2001
DEFINITION	Homo sapiens genomic DNA, chromosome 8q23, clone:KB1670H7.
ACCESSION	AF002848
VERSION	
KEYWORDS	GI:12328451
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

AP002848  
 Homo sapiens genomic DNA, chromosome 8q23, clone:KB1670H7.  
 AF002848  
 GI:12328451  
 Homo sapiens cell\_line:FLEB 14 - 14 DNA, clone\_lib:Keio BAC library  
 clone:KB1670H7.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 183125)  
 Shimizu N. and Asakawa,S.  
 Homo sapiens chromosome clone KB1670H7 on 8q23  
 Published Only in Database (2000) In press  
 2 (bases 1 to 183125)  
 Shimizu N. and Asakawa,S.  
 Direct Submission  
 Submitted (10-AUG-2000) Nobuyoshi Shimizu, Keio University, School  
 of Medicine, Molecular Biology; 35 Shinanomachi, Shinjuku-ku, Tokyo  
 160-8582, Japan (E-mail:nshimizu@med.keio.ac.jp,  
 Tel:-81-3-3351-2370, Fax:81-3-3351-2370)  
 On Jan 19, 2001 this sequence version replaced gi:9798632.  
 Location/Qualifiers  
 1..183125

```

/db_xref="taxon:9606"
/chromosome="8"
/map="8q23"
/clone="KB1670H7"
/cell_line="FLEB 14 - 14"

```



repeat_region	/clone_lib="Kelo BAC library"	repeat_region	complement(20412..21003)
	complement(119..1004)		/evidence-not_experimental
	/rpt_family="L1PA13"	repeat_region	/rpt_family="HERV16"
	complement(5651..5736)		complement(21009..21269)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L2"	repeat_region	/rpt_family="L1PA16"
	complement(5870..7113)		complement(21559..21801)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="HERVL"	repeat_region	/rpt_family="NAL1"
	complement(7215..8911)		22760..22790
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="HERVL"	repeat_region	/rpt_family="(TC)n"
	complement(9824..9923)		22791..22833
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MIR"	repeat_region	/rpt_family="(CA)n"
	10182..10506		22853..22873
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L2"	repeat_region	/rpt_family="AT_rich"
	10507..10798		23766..23825
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="AlusX"	repeat_region	/rpt_family="L2"
	10799..10895		complement(26443..26521)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L2"	repeat_region	/rpt_family="MADE1"
	11776..11956		26702..26766
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MIR"	repeat_region	/rpt_family="MLT1J"
	13212..13232		26767..27179
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="(TC)n"	repeat_region	/rpt_family="MLT2G"
	complement(13752..13836)		27180..27403
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MIR"	repeat_region	/rpt_family="MLT1J"
	complement(13969..14098)		27419..27524
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L2"	repeat_region	/rpt_family="L2"
	complement(14307..14557)		complement(27528..27609)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MIR"	repeat_region	/rpt_family="L2"
	complement(14579..14677)		27830..27862
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L2"	repeat_region	/rpt_family="AT_rich"
	14820..14998		complement(28038..28300)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MLT1H"	repeat_region	/rpt_family="MLT1K"
	complement(15176..15653)		complement(28397..28851)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L1PA2"	repeat_region	/rpt_family="MER70A"
	complement(15666..15940)		29183..29494
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MLT1B"	repeat_region	/rpt_family="AluY"
	complement(15994..16116)		29731..29753
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MLT1B"	repeat_region	/rpt_family="(TTTTC)n"
	16746..16902		complement(30040..30244)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MER5A"	repeat_region	/rpt_family="MIR"
	17054..17186		complement(30280..30581)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MLT2E"	repeat_region	/rpt_family="AluJb"
	17187..17595		30861..30888
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MLT1A2"	repeat_region	/rpt_family="AT_rich"
	17596..17878		31384..31512
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="MLT2E"	repeat_region	/rpt_family="L2"
	complement(17884..18823)		32261..32348
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="L1"	repeat_region	/rpt_family="MIR"
	complement(19188..19990)		complement(33414..33633)
	/evidence-not_experimental		/evidence-not_experimental
	/rpt_family="HERV16"	repeat_region	complement(33782..34282)

```

/evidence-not_experimental
/rpt_family="LIP2"
repeat_region complement(34414..34574)
/evidence-not_experimental
/rpt_family="AluY"
repeat_region 35473..35652
/evidence-not_experimental
/rpt_family="MIR"
repeat_region 35689..36037
/evidence-not_experimental
/rpt_family="MLT1"
repeat_region 36911..37100
/evidence-not_experimental
/rpt_family="MIR"
repeat_region complement(37413..37458)
/evidence-not_experimental
/rpt_family="L2"
repeat_region complement(37459..37482)
/evidence-not_experimental
/rpt_family="MIR"
repeat_region complement(37650..37951)
/evidence-not_experimental
/rpt_family="LIR50"
repeat_region complement(37919..38131)

Alignment Scores:
Pred. No.: 542 Length: 183125
Score: 61.00 Matches: 12
Percent Similarity: 93.338 Conservative: 2
Best Local Similarity: 80.008 Mismatches: 1
Query Match: 81.338 Indels: 0
DB: 9 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AP002848 (1-183125)
QY 1 SerSerSerIleleucysSerProleuProSerleuSerleuSer 15
Db 149928 ACACATCAATTCGTGCTCACCTCCCAACCTCAGCTCTCT 149884

RESULT 5
AC084114/C 189127 bp DNA linear PRI 01-FEB-2002
LOCUS AC084114 Homo sapiens chromosome 8, clone RP11-1059L18, complete sequence.
DEFINITION AC084114
VERSION AC084114.8 GI:18464106
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 189127)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 8, clone RP11-1059L18
2 (bases 1 to 189127)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouhagalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Lacroque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Marcuis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Menus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Olivat,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sounez,C., Spencer,B., Stange-Thoman,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tirrell,A., Travers,M., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,

```

```

Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,C., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (13-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 189127)
REFERENCE
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhagalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Chapel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamati,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marcuis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Menus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Orbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Rietta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thoman,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,C.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 1, 2002 this sequence version replaced g1:18450000.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10840
Center clone name: 1059_L_18
-----
Location/Qualifiers
1. 189127
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-1059L18"
/clone_lib="RP11-11 Human Male BAC"
complement(14..319)
/rpt_family="MLT1"
repeat_region 113..117
/note="<30 qual SINGL region."
167..171
/note="<30 qual SINGL region."
1550..2067
/rpt_family="MLT1H1"
repeat_region 2213..2400
/rpt_family="MIR3"
complement(3205..3249)
/rpt_family="L2"
repeat_region 3250..3563
/rpt_family="AluX"
complement(3564..3608)
/rpt_family="L2"
repeat_region complement(3927..4359)
/rpt_family="LIME1"
repeat_region 4481..4806
/rpt_family="L2"
repeat_region 4844..4974

```

repeat_region	/rpt_family="L1MB8" 5477..5791	
repeat_region	/rpt_family="A1USx" 6550..6590	
repeat_region	/rpt_family="TGCn" complement(6605..6831)	
repeat_region	/rpt_family="MLTIE2" complement(6682..7762)	
repeat_region	/rpt_family="L1MC4" complement(7763..7980)	
repeat_region	/rpt_family="L1MA1" complement(7981..8070)	
repeat_region	/rpt_family="L1MC4" complement(8101..8163)	
repeat_region	/rpt_family="L1MC2" complement(8159..8340)	
repeat_region	/rpt_family="L1MC3" complement(8341..8640)	
repeat_region	/rpt_family="A1USg" complement(8641..8967)	
repeat_region	/rpt_family="L1MC3" 8968..8993	
repeat_region	/rpt_family="(TTTA)n" complement(8994..9235)	
repeat_region	/rpt_family="L1MC3" 9233..9440	
repeat_region	/rpt_family="L1MC4a" complement(9444..9695)	
repeat_region	/rpt_family="L1MBC" complement(9755..9933)	
repeat_region	/rpt_family="L1MBC" 10007..10042	
repeat_region	/rpt_family="AT_rich" complement(110411..10663)	
repeat_region	/rpt_family="MLT1K" complement(110703..11218)	
repeat_region	/rpt_family="L2" complement(11251..11499)	
repeat_region	/rpt_family="MLT1K" complement(112557..12621)	
repeat_region	/rpt_family="L2" 12819..13178	
repeat_region	/rpt_family="L2" 13372..13379	
repeat_region	/rpt_family="L2" complement(133423..13599)	
repeat_region	/rpt_family="L2" complement(14138..15236)	
repeat_region	/rpt_family="MLT1U" complement(15831..15971)	
repeat_region	/rpt_family="LTR27" complement(16004..16578)	
repeat_region	/rpt_family="LTR27" 16590..17021	
repeat_region	/rpt_family="MLT1C" 17022..17479	
repeat_region	/rpt_family="L2" 17745..17871	
repeat_region	/rpt_family="L2" 18053..18211	
repeat_region	/rpt_family="MIR" 19166..19187	
repeat_region	/rpt_family="(T)n" complement(19989..20181)	
repeat_region	/rpt_family="MIR" 20234..20579	
repeat_region	/rpt_family="L1M4" 21771..21911	
repeat_region	/rpt_family="MIR3" 22077..22098	
repeat_region	/rpt_family="AT_rich" complement(22283..22538)	
repeat_region	/rpt_family="MLT1I" 22538..22538	

```
repeat_region      complement(22792..22944)
                    /rpt_family="MIR"
                    23103..23250
repeat_region      /rpt_family="MIR"
                    23394..23337
repeat_region      /rpt_family="(CAT)n"
                    23737..24055
repeat_region      /rpt_family="L2"
                    24164..24244
repeat_region      /rpt_family="LRR33A"
                    24434..24517
repeat_region      /rpt_family="L2"
                    complement(24797..25032)
repeat_region      /rpt_family="L2"
                    complement(27098..27504)
repeat_region      /rpt_family="MIR"
                    27711..27760
repeat_region      /rpt_family="(ATATAG)n"
                    complement(28712..28808)
repeat_region      /rpt_family="L2"
                    complement(28810..29053)
repeat_region      /rpt_family="LNR3"
                    29164..29244
```

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC084114 (1-189127)

```

QY      1 SerSerSerIleLeuCysSerProLeuProSerLeuSerLeuSer 15
      :: |||||
DB    36618 ACACATCATTCCTGCTCACCTCTCCCAACCTCAGCCCTCTCT 36574

```

RESULT 6  
AC079196

LOCUS	195477 bp	DNA	linear	HTG 18-NOV-2000
AC079196				
DEFINITION	Homo sapiens chromosome 8 clone RP11-480P5 map 8, WORKING DRAFT			
SEQUENCE	23 unordered pieces.			

VERSION AC079196.2 GI:11225390  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT  
SOURCE NCBI

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranita; Vertebrata; Euteleostomi;  
Mammalia; Proboscidea; Eulipotyphla; Carnivora; Canidae; Canis; Canis lupus

REFERENCE  
1 (bases 1 to 195477)  
AUTHORS  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE  
Homo sapiens chromosome 9 short arm p11-4000

JOURNAL  
 REFERENCES  
 2 (bases 1 to 195477)  
 unpublished  
 Birren B     Linton J     Nuchbaum C     Landor E     Abraham H     Allen N

Anderson, S., Baird, N., Baschian, V., Bead, F., Boguslavsky, L.,  
Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,  
Chenhal, V., Collamore, M., Collins, S., Collamore, A., Cooke, P.

Denizlihan, N., Dewail, N., Diaz, O. S., Douge, S., Felleiter, F., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Glinde, S., Goyette, M., Graham, L., Grand-Pierre, N., Haros, R., Heaford, A., Horton, J.,

Aliev, I., Johnson, R., Jones, C., Kamin, Y., Natsalalas, N., Panocque, R.,  
Lamazares, R., Landers, T., Lehocky, J., Levine, R., Lieu, C., Liu, G.,  
Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.

O'Donnell, P., O'Neil, D., Olivares, T. M., Oliver, J., Peterson, K., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., McInerney, S. A., Melchior, J., Meneses, L., Minerva, T., Mwangi, W.,

Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Souarez, C., Spencer, B., Strange-Thomann, N., Stoltanovic, N.

Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zahoun, J.,

TITLE  
JOURNAL  
COMMENT

Zimmer, A. and Zody, M.  
Submitted (23-ANG-2000) Whitehead Institute/MIT Center for Genome  
Research, 330 Charles Street, Cambridge, MA 02141, USA  
On Nov 18, 2000 this sequence version replaced g1:9886032.  
All repeats were identified using RepeatMasker:  
Shtlt, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L10608

Center clone name: 480.F.5

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 180424 bases at least Q40  
Consensus quality: 188415 bases at least Q30  
Consensus quality: 191442 bases at least Q20  
Insert size: 198000; agarose-1p  
Insert size: 193277; sum-of-coverage  
Quality coverage: 4.2 in Q20 bases; agarose-1p  
Quality coverage: 4.3 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 23 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
776 875: contig of 775 bp in length  
776 875: gap of 100 bp  
876 1876: contig of 1001 bp in length  
1877 1976: gap of 100 bp  
1977 58986: contig of 57010 bp in length  
58987 59086: gap of 100 bp  
59087 60289: contig of 1203 bp in length  
60290 60389: gap of 100 bp  
60390 61513: contig of 1124 bp in length  
61514 61613: gap of 100 bp  
61614 62903: contig of 1290 bp in length  
62904 63003: gap of 100 bp  
63004 64788: contig of 1785 bp in length  
64789 64888: gap of 100 bp  
64889 67987: contig of 3099 bp in length  
67988 68087: gap of 100 bp  
68088 71902: contig of 3815 bp in length  
71903 72002: gap of 100 bp  
72003 74471: contig of 2469 bp in length  
74472 74571: gap of 100 bp  
74572 77764: contig of 3193 bp in length  
77765 77864: gap of 100 bp  
77865 81116: contig of 3252 bp in length  
81117 81216: gap of 100 bp  
81217 85155: contig of 3939 bp in length  
85156 85255: gap of 100 bp  
85256 89134: contig of 3879 bp in length  
89135 89234: gap of 100 bp  
89235 96530: contig of 7296 bp in length  
96531 96630: gap of 100 bp  
96631 104817: contig of 8187 bp in length  
104818 104917: gap of 100 bp  
104918 118601: contig of 13684 bp in length  
118602 118701: gap of 100 bp  
118702 129776: contig of 11075 bp in length  
129777 129876: gap of 100 bp  
129877 140248: contig of 10372 bp in length

# FEATURES

140249 140348: gap of 100 bp  
140349 150907: contig of 10559 bp in length  
150908 151007: gap of 100 bp  
151008 168203: contig of 18196 bp in length  
168204 169303: gap of 100 bp  
169304 194921: contig of 25618 bp in length  
194922 195021: gap of 100 bp  
195022 195477: contig of 456 bp in length.  
Location/Qualifiers

1. 195477  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-480P5"  
/clone\_11b="RP11 Human Male BAC"  
1. 775  
/note="assembly-fragment  
clone\_end:SP6  
vector\_side:left"  
876..1876  
/note="assembly-fragment"  
1977..58986  
/note="assembly-fragment"  
59087..60289  
/note="assembly-fragment"  
60390..61513  
/note="assembly-fragment"  
61614..62903  
/note="assembly-fragment"  
63004..64788  
/note="assembly-fragment"  
64889..67987  
/note="assembly-fragment"  
68088..71902  
/note="assembly-fragment"  
72003..74471  
/note="assembly-fragment"  
74572..77764  
/note="assembly-fragment"  
77865..81116  
/note="assembly-fragment"  
81217..85155  
/note="assembly-fragment"  
85256..89134  
/note="assembly-fragment"  
89235..96530  
/note="assembly-fragment"  
96631..104817  
/note="assembly-fragment"  
104918..118601  
/note="assembly-fragment"  
118702..129776  
/note="assembly-fragment"  
129777..140248  
/note="assembly-fragment"  
140349..150907  
/note="assembly-fragment"  
151008..169303  
/note="assembly-fragment"  
169304..194921  
/note="assembly-fragment"  
195022..195477  
/note="assembly-fragment  
clone\_end:T7  
vector\_side:right"

BASE COUNT 58603 a 36972 c 36289 g 61402 t 2211 others  
ORIGIN

Alignment Scores: 579 Length: 195477  
Pred. No.: 61.00 Matches: 12  
Score: 93.33% Conservative: 2  
Percent Similarity:

Best Local Similarity: 80.00%  
Query Match: 81.33%  
DB: 2  
Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AC079196 (1-195477)

0y 1 SerSerSerIleLeuGysSerProLeuProSerLeuSerLeuSer 15

Db 46364 ACACAAATCAATTCGTCTCCTCACCTCCCAACCTCACCTCTCT 46408

RESULT 7

AC109887 91582 bp DNA linear HTG 13-JUL-2002

LOCUS Rattus norvegicus clone CH230-312G3, \*\*\* SEQUENCING IN PROGRESS

DEFINITION \*\*\* 49 unordered pieces.

AC109887.3 GI:21738183

HTG: HTGS-PHASE1.

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

Rattus.

1 (bases 1 to 91582)

Murphy,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,

Alabrooks,S.L., Amarantunga,H.C., Are,J.R., Ayele,M., Banks,T.,

Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,

Bunyah,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Dem,A.L., Ding,X., Dinh,H.H.,

Douthaite,K.J., Drepper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorell,J.D.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

Honsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovai,C.,

Kravic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Lichtenberg,O., Lieu,C., Liu,J., Liu,W., Louissege,H.,

Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Masey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

Nguyen,N., Nickerson,E., Nwokoko,S., Ogih,M., Okunolu,G.,

Orangunye,N., Oyedero,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,

Scheier,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,

Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Syatek,A., Tabor,P., Tameris,A., Tameris,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Umanu,K., Vasquez,U., Vera,V., Villalón,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,

Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K.,

Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 91582)

Worley,K.C.

Direct Submission

Submitted (08-FEB-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 91582)

AUTHORS  
TITLE  
JOURNAL

COMMENT

Worley,K.C.  
Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18860208.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GOKI

Center clone name: CH230-312G3

Summary Statistics

Sequencing vector: Plasmid:

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 62571 bases at least Q40

Consensus quality: 66037 bases at least Q30

Consensus quality: 69135 bases at least Q20

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 49 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1503 1502: contig of 1502 bp in length

1803 2721: contig of 1119 bp in length

2722 2821: gap of unknown length

2822 3963: contig of 1142 bp in length

3964 4063: gap of unknown length

4064 5105: contig of 1042 bp in length

5106 5205: gap of unknown length

5206 6577: contig of 1372 bp in length

6578 6677: gap of unknown length

6678 7902: contig of 1125 bp in length

7903 8946: contig of 1044 bp in length

8947 9046: gap of unknown length

9047 10117: contig of 1071 bp in length

10118 10217: gap of unknown length

10218 11479: contig of 1262 bp in length

11480 11579: gap of unknown length

11580 12844: contig of 1265 bp in length

12845 12944: gap of unknown length

12945 14950: contig of 2006 bp in length

14951 15050: gap of unknown length

15051 16642: contig of 1592 bp in length

16643 16742: gap of unknown length

16743 17828: contig of 1086 bp in length

17829 17928: gap of unknown length

17929 19727: contig of 1799 bp in length

19728 21363: gap of unknown length

21364 21463: gap of 1536 bp in length

21464 22652: contig of 1189 bp in length

22653 22752: gap of unknown length

22753 23929: contig of 1177 bp in length

23930 24029: gap of unknown length

24030 25578: contig of 1549 bp in length

25579 25678: gap of unknown length

25679 27062: contig of 1284 bp in length

27063 28472: gap of unknown length

28473 28572: contig of 1410 bp in length

28573 30264: gap of unknown length

30264: contig of 1692 bp in length

\* 30265 30364: gap of unknown length  
\* 30365 31464: contig of 1100 bp in length  
\* 31465 31564: gap of unknown length  
\* 31565 33215: contig of 1651 bp in length  
\* 33216 33315: gap of unknown length  
\* 33316 34890: contig of 1575 bp in length  
\* 34891 34990: gap of unknown length  
\* 34991 36624: contig of 1634 bp in length  
\* 36625 36724: gap of unknown length  
\* 36725 38529: contig of 1805 bp in length  
\* 38530 38630: gap of unknown length  
\* 38630 40251: contig of 1622 bp in length  
\* 40252 40351: gap of unknown length  
\* 40352 42287: contig of 1936 bp in length  
\* 42288 42387: gap of unknown length  
\* 42388 43463: contig of 1076 bp in length  
\* 43464 43564: gap of unknown length  
\* 43564 44922: contig of 1355 bp in length  
\* 44922 45022: gap of unknown length  
\* 45023 47349: contig of 2327 bp in length  
\* 47350 47449: gap of unknown length  
\* 47450 48918: contig of 1469 bp in length  
\* 48919 49018: gap of unknown length  
\* 49019 51017: contig of 1999 bp in length  
\* 51018 51117: gap of unknown length  
\* 51118 52930: contig of 1813 bp in length  
\* 52931 53031: gap of unknown length  
\* 53031 54539: contig of 1509 bp in length  
\* 54540 54639: gap of unknown length  
\* 54640 57214: contig of 2575 bp in length  
\* 57215 57314: gap of unknown length  
\* 57315 59204: contig of 1890 bp in length  
\* 59205 59304: gap of unknown length  
\* 59305 60829: contig of 1525 bp in length  
\* 60830 60929: gap of unknown length  
\* 60930 63121: contig of 2193 bp in length  
\* 63122 63221: gap of unknown length  
\* 63222 65614: contig of 2393 bp in length  
\* 65615 65714: gap of unknown length  
\* 65715 67972: contig of 2258 bp in length  
\* 67973 68072: gap of unknown length  
\* 68073 69880: contig of 1808 bp in length  
\* 69881 69980: gap of unknown length  
\* 69981 73676: contig of 3656 bp in length  
\* 73677 73776: gap of unknown length  
\* 73777 76126: contig of 2350 bp in length  
\* 76127 76226: gap of unknown length  
\* 76227 78822: contig of 2596 bp in length  
\* 78823 78922: gap of unknown length  
\* 78923 82017: contig of 3095 bp in length  
\* 82018 82117: gap of unknown length  
\* 82118 84115: contig of 1998 bp in length  
\* 84116 84215: gap of unknown length  
\* 84216 87398: contig of 3183 bp in length  
\* 87399 87498: gap of unknown length  
\* 87499 91582: contig of 4084 bp in length.  
\* Location/Qualifiers  
\* 1. 91582  
\* /organism="Rattus norvegicus"  
\* /db\_xref="taxon:10116"  
\* /clone="CH230-31263"

BASE COUNT 24682 a 18454 c 18294 g 25285 t 4867 others  
ORIGIN

Alignment Scores:  
Pred. No.: 387  
Score: 60.00  
Percent Similarity: 92.868  
Best Local Similarity: 85.714  
Query Match: 80.004  
DB: 2

Length: 91582  
Matches: 12  
Conservative: 1  
Mismatch: 1  
Indels: 0  
Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC109887 (1-91582)

Oy 2 SerSer1leucCySerProleuProserleuSerleuSer 15  
Db 65489 TCCAGCATCTCTGTTCTCATTTCTTCTACACGTCTCC 65530

RESULT 8  
AC095732/c 169845 bp DNA linear HTG 11-JUL-2002  
LOCUS Rattus norvegicus clone CH230-9A18, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
DEFINITION 67 unordered pieces.  
ACCESSION AC095732  
VERSION AC095732.4 GI:21722742  
KEYWORDS HTG, HTGS, PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus

REFERENCE  
AUTHORS 1 (bases 1 to 169845)  
Muzny,D.M., Adams,C., Adio-Oduola,B., All-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barberia,J., Benton,J., Bimge,K., Blankenburg,K., Bonini,D.,  
Bouck,J., Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Devalla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durkin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Fallis,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gablis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homst,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssoen,E., Kelly,S., Khan,U., King,L., Koryah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lileu,C., Liu,J., Liu,W., Louisedge,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Mattheis,E.,  
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabab,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nockenwo,S., Oguh,M., Okunou,G.,  
Oragunye,N., Oyiedu,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G.,  
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,  
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Teliford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasequez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
Williams,G., Williamson,A., Wlecczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 169845)  
AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (17-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 169845)  
REFERENCE Direct Submission  
AUTHORS Worley,K.C.  
TITLE Direct Submission  
JOURNAL Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Jul 10, 2002 this sequence version replaced g1:20975871.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GDEZ

Center clone name: CH230-9A18

----- Summary Statistics

Sequencing vector: Plasmid:

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 104990 bases at least Q40

Consensus quality: 112511 bases at least Q30

Consensus quality: 119479 bases at least Q20

-----

\* NOTE: Estimated insert size may differ from sequence length

\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 67 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

1 1038: contig of 1038 bp in length  
\* 1039 1138: gap of unknown length  
\* 1139 2221: contig of 1083 bp in length  
\* 2222 2321: gap of unknown length  
\* 2322 3547: contig of 1226 bp in length  
\* 3548 3647: gap of unknown length  
\* 3648 4983: contig of 1336 bp in length  
\* 4984 5083: gap of unknown length  
\* 5084 6222: contig of 1139 bp in length  
\* 6223 6322: gap of unknown length  
\* 6323 7585: contig of 1263 bp in length  
\* 7586 7685: gap of unknown length  
\* 7686 8858: contig of 1173 bp in length  
\* 8859 8958: gap of unknown length  
\* 8959 10012: contig of 1054 bp in length  
\* 10013 11455: contig of 1343 bp in length  
\* 11456 11555: gap of unknown length  
\* 11556 13017: contig of 1462 bp in length  
\* 13018 13117: gap of unknown length  
\* 13118 14720: contig of 1603 bp in length  
\* 14721 14820: gap of unknown length  
\* 14821 16158: contig of 1338 bp in length  
\* 16159 16258: gap of unknown length  
\* 16259 17750: contig of 1492 bp in length  
\* 17751 17850: gap of unknown length  
\* 17851 19465: contig of 1615 bp in length  
\* 19466 19565: gap of unknown length  
\* 19566 20786: contig of 1221 bp in length  
\* 20787 20886: gap of unknown length  
\* 20887 22283: contig of 1397 bp in length  
\* 22284 22383: gap of unknown length  
\* 22384 23550: contig of 1167 bp in length  
\* 23551 23650: gap of unknown length  
\* 23651 25247: contig of 1597 bp in length  
\* 25248 25347: gap of unknown length  
\* 25348 26739: contig of 1392 bp in length  
\* 26740 26839: gap of unknown length  
\* 26840 28297: contig of 1458 bp in length  
\* 28298 28397: gap of unknown length  
\* 28398 30539: contig of 2142 bp in length  
\* 30540 30639: gap of unknown length  
\* 30640 31984: contig of 1345 bp in length  
\* 31985 32084: gap of unknown length  
\* 32085 33203: contig of 1119 bp in length  
\* 33204 33303: gap of unknown length

\* 33304 35105: contig of 1802 bp in length  
\* 35106 35205: gap of unknown length  
\* 35206 36554: contig of 1349 bp in length  
\* 36555 36654: gap of unknown length  
\* 36655 39009: contig of 2355 bp in length  
\* 39010 39109: gap of unknown length  
\* 39110 40559: contig of 1450 bp in length  
\* 40560 40659: gap of unknown length  
\* 40660 41768: contig of 1109 bp in length  
\* 41769 41868: gap of unknown length  
\* 41869 44085: contig of 2217 bp in length  
\* 44086 44185: gap of unknown length  
\* 44186 46351: contig of 2166 bp in length  
\* 46352 46451: gap of unknown length  
\* 46452 48980: contig of 2529 bp in length  
\* 48981 49080: gap of unknown length  
\* 49081 51093: contig of 2013 bp in length  
\* 51094 51193: gap of unknown length  
\* 51194 52710: contig of 1517 bp in length  
\* 52711 52810: gap of unknown length  
\* 52811 55632: contig of 2822 bp in length  
\* 55633 55732: gap of unknown length  
\* 55733 58525: contig of 2793 bp in length  
\* 58526 58625: gap of unknown length  
\* 58626 61423: contig of 2798 bp in length  
\* 61424 61523: gap of unknown length  
\* 61524 63441: contig of 1918 bp in length  
\* 63442 63541: gap of unknown length  
\* 63542 66215: contig of 2674 bp in length  
\* 66216 66315: gap of unknown length  
\* 66316 68159: contig of 1844 bp in length  
\* 68160 68259: gap of unknown length  
\* 68260 70262: contig of 2003 bp in length  
\* 70263 70362: gap of unknown length  
\* 70363 71758: contig of 1396 bp in length  
\* 71759 71858: gap of unknown length  
\* 71859 73318: contig of 1460 bp in length  
\* 73319 73418: gap of unknown length  
\* 73419 75122: contig of 1704 bp in length  
\* 75123 75222: gap of unknown length  
\* 75223 78668: contig of 3446 bp in length  
\* 78669 78768: gap of unknown length  
\* 78769 81040: contig of 2272 bp in length  
\* 81041 81140: gap of unknown length  
\* 81141 83306: contig of 2166 bp in length  
\* 83307 83406: gap of unknown length  
\* 83407 86465: contig of 3059 bp in length  
\* 86466 86565: gap of unknown length  
\* 86566 89067: contig of 2502 bp in length  
\* 89068 89167: gap of unknown length  
\* 89168 92835: contig of 3668 bp in length  
\* 92836 92935: gap of unknown length  
\* 92936 95306: contig of 2571 bp in length  
\* 95307 95606: gap of unknown length  
\* 95607 98304: contig of 2698 bp in length  
\* 98305 98404: gap of unknown length  
\* 98405 101380: contig of 2976 bp in length  
\* 101381 101480: gap of unknown length  
\* 101481 105478: contig of 3998 bp in length  
\* 105479 105578: gap of unknown length  
\* 105579 108776: contig of 3198 bp in length

## Alignment Scores:

Pred. No.: 2.07e+03 Length: 169845  
Score: 57.00 Matches: 13  
Percent Similarity: 86.67% Conservative: 0  
Best Local Similarity: 86.67% Mismatches: 2  
Query Match: 76.00% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AC095732 (1-169845)

Qy 1 SerSerSerIleuCySerProLeuProSerLeuSerLeuSer 15







```

Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDOO
Center clone name: CH230-9122
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 44880 bases at least Q40
Consensus quality: 54834 bases at least Q30
Consensus quality: 62109 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1024: contig of 1024 bp in length
* 1025 1124: gap of unknown length
* 1125 2135: contig of 1011 bp in length
* 2136 2235: gap of unknown length
* 2236 3600: contig of 1365 bp in length
* 3601 3700: gap of unknown length
* 3701 4797: contig of 1097 bp in length
* 4798 4898: gap of unknown length
* 4898 5991: contig of 1094 bp in length
* 5992 6092: gap of unknown length
* 6092 7247: contig of 1155 bp in length
* 7247 7347: gap of unknown length
* 7347 8487: contig of 1141 bp in length
* 8488 8587: gap of unknown length
* 8588 9639: contig of 1052 bp in length
* 9640 9739: gap of unknown length
* 9740 11090: contig of 1351 bp in length
* 11091 11190: gap of unknown length
* 11191 12196: contig of 1006 bp in length
* 12197 12396: gap of unknown length
* 12397 13886: contig of 1590 bp in length
* 13887 13986: gap of unknown length
* 13987 15685: contig of 1699 bp in length
* 15686 15785: gap of unknown length
* 15786 17641: contig of 1856 bp in length
* 17642 17741: gap of unknown length
* 17741 20288: contig of 2547 bp in length
* 20289 20388: gap of unknown length
* 20389 21497: contig of 1109 bp in length
* 21498 21597: gap of unknown length
* 21598 22668: contig of 1071 bp in length
* 22669 22768: gap of unknown length
* 22769 24455: contig of 1687 bp in length
* 24456 24555: gap of unknown length
* 24556 26278: contig of 1723 bp in length
* 26279 26378: gap of unknown length
* 26379 27628: contig of 1250 bp in length
* 27629 27728: gap of unknown length
* 27729 29350: contig of 1522 bp in length
* 29351 29350: gap of unknown length
* 29351 30860: contig of 1510 bp in length
* 30861 30960: gap of unknown length
* 30961 32051: contig of 1091 bp in length
* 32052 32151: gap of unknown length
* 32152 34551: contig of 2400 bp in length
* 34552 34651: gap of unknown length
* 34652 35991: contig of 1340 bp in length
* 35992 36091: gap of unknown length
* 36092 37109: contig of 1018 bp in length
* 37110 37209: gap of unknown length

```

```

* 37210 38947: contig of 1738 bp in length
* 38948 39047: gap of unknown length
* 39048 40168: contig of 1121 bp in length
* 40169 40268: gap of unknown length
* 40269 41542: contig of 1274 bp in length
* 41543 41642: gap of unknown length
* 41643 43437: contig of 1795 bp in length
* 43438 43537: gap of unknown length
* 43538 45119: contig of 1582 bp in length
* 45120 45219: gap of unknown length
* 45220 46865: contig of 1646 bp in length
* 46866 46965: gap of unknown length
* 46966 48123: contig of 1158 bp in length
* 48124 48224: gap of unknown length
* 48224 49481: contig of 1288 bp in length
* 49482 49581: gap of unknown length
* 49582 50995: contig of 1414 bp in length
* 50996 51095: gap of unknown length
* 51096 52439: contig of 1344 bp in length
* 52440 52539: gap of unknown length
* 52540 54279: contig of 1740 bp in length
* 54280 54379: gap of unknown length
* 54380 56635: contig of 2256 bp in length
* 56636 56735: gap of unknown length
* 56736 58377: contig of 1642 bp in length
* 58378 58477: gap of unknown length
* 58478 59989: contig of 1512 bp in length
* 59990 60089: gap of unknown length
* 60090 61584: contig of 1495 bp in length
* 61585 61684: gap of unknown length
* 61685 62912: contig of 1228 bp in length
* 62913 63012: gap of unknown length
* 63013 64579: contig of 1567 bp in length
* 64580 64679: gap of unknown length
* 64680 66134: contig of 1455 bp in length
* 66135 66234: gap of unknown length
* 66235 68459: contig of 2225 bp in length
* 68460 68559: gap of unknown length
* 68560 69792: contig of 1233 bp in length
* 69793 69892: gap of unknown length
* 69893 71547: contig of 1655 bp in length
* 71548 71647: gap of unknown length
* 71648 73376: contig of 1729 bp in length
* 73377 73476: gap of unknown length
* 73477 76099: contig of 2623 bp in length
* 76100 76199: gap of unknown length
* 76200 78646: contig of 2447 bp in length
* 78647 78746: gap of unknown length
* 78747 80040: contig of 1294 bp in length
* 80041 80140: gap of unknown length
* 80141 82556: contig of 2416 bp in length
* 82557 82656: gap of unknown length
* 82657 84373: contig of 1617 bp in length
* 84374 84373: gap of unknown length
* 84374 86987: contig of 2614 bp in length
* 86988 87087: gap of unknown length
* 87088 89267: contig of 2180 bp in length

```

```

Alignment Scores:
Pred. No.: 1 82e+03 Length: 104525
Score: 56.00 Matches: 11
Percent Similarity: 86.67% Conservative: 2
Best Local Similarity: 73.33% Mismatches: 2
Query Match: 74.67% Indels: 0
DB: 2 Gaps: 0

```

```

US-10-075-846-4_COPY_192_207 (1-16) x AC0095220 (1-104525)

```

```

QY 1 SerSerSeriLeuCySerProLeuProSerLeuSerLeuSer 15
||||| ::||| ||||| ||||| ||||| ||||| |||||
DB 43411 TCATCGTTCCTCTCTGCTCCCTCTCTTCAATCACTGCTCC 43367
RESULT 11

```

AC102866  
LOCUS AC102866 172555 bp DNA linear HTG 20-AUG-2002  
DEFINITION Mus musculus clone RP24-576L9, WORKING DRAFT SEQUENCE, 12 unordered  
pieces.  
AC102866  
VERSION AC102866.2 GI:22325064  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HMGS-FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 172555)  
AUTHORS Birren, B., Nusbaum, C. and Lander, E.  
TITLE Mus musculus, clone RP24-576L9  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 172555)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,  
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, M., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K.,  
Lamarez, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Olivier, J., Peterson, K., Punukhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, N., Schupack, R.,  
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Strauss, K., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
JOURNAL Direct Submission  
TITLE Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA  
AUTHORS 3 (bases 1 to 172555)  
Birren, B., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepe, Y., Collymore, A.,  
Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Olivier, J., Peterson, K.,  
Punukhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, N., Schupack, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,  
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.  
JOURNAL Direct Submission  
TITLE Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome  
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Aug 20, 2002 this sequence version replaced g1:17060976.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information

Center project name: L19975  
Center clone name: 576\_L-9  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 169439 bases at least Q40  
Consensus quality: 170530 bases at least Q30  
Consensus quality: 170959 bases at least Q20  
Insert size: 157000; agarose-fp  
Insert size: 171455; sum-of-contigs  
Quality coverage: 12.8 in Q20 bases; agarose-fp  
Quality coverage: 11.7 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
635 734: gap of 100 bp in length  
735 1427: contig of 693 bp in length  
1428 1527: gap of 100 bp  
1528 2188: contig of 661 bp in length  
2189 2288: gap of 100 bp  
2289 2980: contig of 692 bp in length  
2981 3080: gap of 100 bp  
3081 3718: contig of 638 bp in length  
3719 3818: gap of 100 bp  
3819 4462: contig of 644 bp in length  
4463 4562: gap of 100 bp  
4563 5222: contig of 660 bp in length  
5223 5322: gap of 100 bp  
5323 6387: contig of 1065 bp in length  
6388 6487: gap of 100 bp  
6488 8217: contig of 1730 bp in length  
8218 8317: gap of 100 bp  
8318 9778: contig of 1461 bp in length  
9779 9878: gap of 100 bp  
9879 89823: contig of 79845 bp in length  
89824 172555: contig of 82732 bp in length.  
-----  
FEATURES  
source  
1. 172555  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone.lib="RP24-576L9"  
/clone.lib="RP24-576L9"  
1. 634  
/note="assembly-fragment"  
735. 1427  
/note="assembly-fragment"  
1528. 2188  
/note="assembly-fragment"  
2289. 2980  
/note="assembly-fragment"  
3081. 3718  
/note="assembly-fragment"  
3819. 4462  
/note="assembly-fragment"  
4563. 5222  
/note="assembly-fragment"  
5323. 6387  
/note="assembly-fragment"  
6488. 8217  
/note="assembly-fragment"  
8318. 9778  
/note="assembly-fragment"  
9879. 89723  
/note="assembly-fragment"

```

misc_feature      89824..172555
                  /note="assembly_fragment"
BASE COUNT       53892 a 31196 c 31503 g 54857 t 1107 others
ORIGIN
Alignment Scores:
Pred. No.:       3e+03      Length:      172555
Score:           56.00      Matches:      10
Percent Similarity: 100.00%  Conservative: 3
Best Local Similarity: 76.92%  Mismatches: 0
Query Match:     74.67%     Indels:      0
DB:              2         Gaps:         0

US-10-075-846-4_COPY_192_207 (1-16) x AC102866 (1-172555)
QY
4 ILLeucySerProLeuProSerLeuSerLeuSerVal 16
Db      52101 TTGCTTGCCTCCCTCCCTCCGTCATGTCCTCTCTCT 52139

RESULT 12
GG126MRP      598 bp      mRNA      linear      VRT 28-APR-1992
LOCUS         GG126MRP      598 bp      mRNA      linear      VRT 28-APR-1992
DEFINITION   G.gallus 126 MRP mRNA.
ACCESSION    X61200
VERSION      X61200.1 GI:62998
KEYWORDS
SOURCE
ORGANISM     Chicken.
              Gallus gallus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae;
              Phasianinae; Gallus.
REFERENCE
AUTHORS      Graf,T.H.
TITLE        Direct Submission
JOURNAL      Submitted (01-AUG-1991) T.H. Graf, EMBL, Meyerhofstr. 1, 6900
              Heidelberg, FRG
2 (bases 1 to 598)
AUTHORS      Nakano,T. and Graf,T.
TITLE        Identification of genes differentially expressed in two types of
              v-myb-transformed avian myelomonocytic cells
JOURNAL      Oncogene 7 (3), 527-534 (1992)
MEDLINE      92195690
PUBMED
FEATURES
source
Location/Qualifiers
1..598
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="126"
/tissue_lib="AMV(UL106) promyelocytes"
BASE COUNT     158 a 198 c 129 g 113 t
ORIGIN
Alignment Scores:
Pred. No.:       15.1      Length:      598
Score:           55.00      Matches:      11
Percent Similarity: 81.25%  Conservative: 2
Best Local Similarity: 68.75%  Mismatches: 3
Query Match:     73.33%     Indels:      0
DB:              5         Gaps:         0

US-10-075-846-4_COPY_192_207 (1-16) x GG126MRP (1-598)
QY
1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
Db      482 TCCCTCCACACCCCTGCTCCCTCCACACCCCTCCCTCCATC 529

RESULT 13
AC118327      139754 bp      DNA      linear      HTG 14-JUL-2002
LOCUS         AC118327      139754 bp      DNA      linear      HTG 14-JUL-2002
DEFINITION   Rattus norvegicus clone CH230-501B21, *** SEQUENCING IN PROGRESS
ACCESSION    AC118327
VERSION      AC118327.2 GI:21747379

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
1 (bases 1 to 139754)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbata,J., Benton,J., Birlewa,M., Brown,M., Brown,M., Bryant,N.P.,
Bouck,J., Bowle,S., Brilleman,K., Burdick,C., Burdick,K.L., Byrd,N.C.,
Bunney,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Dayala,M.L., Davis,C., Davy-Carroll,L., Dedrich,B.,
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotco,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
Homsel,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Luchter,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Medford,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokoko,S., Ogun,M., Okunolu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojokokan,I., Rojokokan,R., Ruiz,S., Savary,G.,
Scherrer,S., Scott,G., Shen,H., Shostakov,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Taber,P., Tameris,A., Tameris,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telirod,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 139754)
Worley,K.C.
Direct Submission
Submitted (15-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 139754)
Worley,K.C.
Direct Submission
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20149387.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GVZL
Center clone name: CH230-501B21
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads

```



Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,  
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Morley, K.,  
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
Meinstock, G. and Gibbs, R.  
Direct Submission  
Unpublished  
JOURNAL  
TITLE  
2 (bases 1 to 164772)  
REFERENCE  
AUTHORS  
Morley, K.C.  
JOURNAL  
TITLE  
Submitted (16-SEP-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 164772)  
Morley, K.C.  
Direct Submission  
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 10, 2002 this sequence version replaced g1:17942274.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GDAM  
Center clone name: CH230-9D7  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye 1000 of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 79000 bases at least Q40  
Consensus quality: 85421 bases at least Q20  
Consensus quality: 90756 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 85 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1297: contig of 1297 bp in length  
\* 1398 1397: gap of unknown length  
\* 1398 2679: contig of 1282 bp in length  
\* 2680 2779: gap of unknown length  
\* 2780 3802: contig of 1023 bp in length  
\* 3803 3902: gap of unknown length  
\* 3903 4951: contig of 1049 bp in length  
\* 4952 5051: gap of unknown length  
\* 5052 6299: contig of 1248 bp in length  
\* 6300 6399: gap of unknown length  
\* 6400 7431: contig of 1032 bp in length  
\* 7432 7531: gap of unknown length  
\* 7532 8548: contig of 1017 bp in length  
\* 8549 8648: gap of unknown length  
\* 8649 9775: contig of 1128 bp in length  
\* 9776 9876: gap of unknown length  
\* 9877 10877: contig of 1001 bp in length  
\* 10878 10977: gap of unknown length  
\* 10978 12067: contig of 1090 bp in length  
\* 12068 12167: gap of unknown length  
\* 12168 13239: contig of 1072 bp in length  
\* 13240 13339: gap of unknown length  
\* 13340 14778: contig of 1439 bp in length  
\* 14779 14878: gap of unknown length  
\* 14879 16415: contig of 1537 bp in length  
\* 16416 16515: gap of unknown length  
\* 16516 17592: contig of 1077 bp in length  
\* 17593 17692: gap of unknown length  
\*  
\* 17693 18970: contig of 1278 bp in length  
\* 18971 19070: gap of unknown length  
\* 19071 20403: contig of 1333 bp in length  
\* 20404 20503: gap of unknown length  
\* 20504 21651: contig of 1148 bp in length  
\* 21652 21751: gap of unknown length  
\* 21752 22887: contig of 1136 bp in length  
\* 22888 22988: gap of unknown length  
\* 22989 24446: contig of 1459 bp in length  
\* 24447 24546: gap of unknown length  
\* 24547 25565: contig of 1019 bp in length  
\* 25566 25665: gap of unknown length  
\* 25666 27036: contig of 1371 bp in length  
\* 27037 27136: gap of unknown length  
\* 27137 28479: contig of 1342 bp in length  
\* 28479 28578: gap of unknown length  
\* 28579 29664: contig of 1086 bp in length  
\* 29665 31068: gap of unknown length  
\* 31069 31168: contig of 1304 bp in length  
\* 31169 32406: gap of unknown length  
\* 32407 32506: contig of 1238 bp in length  
\* 32507 34145: gap of unknown length  
\* 34146 34245: contig of 1639 bp in length  
\* 34246 35811: gap of unknown length  
\* 35812 35911: contig of 1566 bp in length  
\* 35912 37018: gap of unknown length  
\* 37019 37118: contig of 1107 bp in length  
\* 37119 38624: gap of unknown length  
\* 38625 38725: contig of 1506 bp in length  
\* 38726 40409: gap of unknown length  
\* 40410 40509: contig of 1685 bp in length  
\* 40510 42002: gap of unknown length  
\* 42003 42102: contig of 1493 bp in length  
\* 42103 43186: gap of unknown length  
\* 43187 43259: contig of 1056 bp in length  
\* 43260 44262: gap of unknown length  
\* 44263 44362: contig of 1004 bp in length  
\* 44363 45568: gap of unknown length  
\* 45569 47129: contig of 1206 bp in length  
\* 47130 47229: gap of unknown length  
\* 47230 48940: contig of 1461 bp in length  
\* 48941 49040: gap of unknown length  
\* 49041 50571: contig of 1711 bp in length  
\* 50572 50671: gap of unknown length  
\* 50672 52207: contig of 1536 bp in length  
\* 52208 52307: gap of unknown length  
\* 52308 53449: contig of 1142 bp in length  
\* 53450 53549: gap of unknown length  
\* 53550 54869: contig of 1320 bp in length  
\* 54870 54969: gap of unknown length  
\* 54970 56026: contig of 1057 bp in length  
\* 56027 56126: gap of unknown length  
\* 56127 58611: contig of 2485 bp in length  
\* 58612 58712: gap of unknown length  
\* 58713 60042: contig of 1331 bp in length  
\* 60043 60142: gap of unknown length  
\* 60143 61451: contig of 1309 bp in length  
\* 61452 61551: gap of unknown length  
\* 61552 63164: contig of 2063 bp in length  
\* 63165 65157: gap of unknown length  
\* 65158 65257: contig of 1443 bp in length  
\* 65258 66957: gap of unknown length  
\* 66958 67057: contig of 1700 bp in length  
\* 67058 69461: gap of unknown length  
\* 69462 69561: contig of 2404 bp in length  
\* 69562 70953: gap of unknown length  
\* 70954 71053: contig of 1392 bp in length  
\* 71054 72992: gap of unknown length  
\* 72993 73092: contig of 1939 bp in length  
\* 73093 75634: gap of unknown length  
\* 75635 75634: contig of 2542 bp in length



```
* 31993 37290: contig of 5298 bp in length
* 37291 37390: gap of 100 bp
* 37391 45486: contig of 8096 bp in length
* 45487 45586: gap of 100 bp
* 45587 53572: contig of 7986 bp in length
* 53573 53672: gap of 100 bp
* 53673 66319: contig of 12647 bp in length
* 66320 66419: gap of 100 bp
* 66420 80154: contig of 13735 bp in length
* 80155 80254: gap of 100 bp
* 80255 94472: contig of 14218 bp in length
* 94473 94572: gap of 100 bp
* 94573 120577: contig of 26005 bp in length
* 120578 120677: gap of 100 bp
* 120678 147069: contig of 26392 bp in length
* 147070 147169: gap of 100 bp
* 147170 174354: contig of 27185 bp in length.
```

## FEATURES

```
source
1. 174354
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-232G11"
/clone_lib="RPC1-23 Female Mouse BMC"
1. 101
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
misc_feature
202. 1145
/note="assembly-fragment"
1246. 2270
/note="assembly-fragment"
misc_feature
2371. 4066
/note="assembly-fragment"
4167. 5725
/note="assembly-fragment"
misc_feature
5826. 8038
/note="assembly-fragment"
8139. 11089
/note="assembly-fragment"
misc_feature
11190. 13192
/note="assembly-fragment"
13293. 16784
/note="assembly-fragment"
misc_feature
16885. 20475
/note="assembly-fragment"
20576. 27499
/note="assembly-fragment"
misc_feature
27600. 31892
/note="assembly-fragment"
31993. 37290
/note="assembly-fragment"
misc_feature
37391. 45486
/note="assembly-fragment"
45587. 53572
/note="assembly-fragment"
misc_feature
53673. 66319
/note="assembly-fragment"
66420. 80154
/note="assembly-fragment"
misc_feature
80255. 94472
/note="assembly-fragment"
94573. 120577
/note="assembly-fragment"
120678. 147069
/note="assembly-fragment"
147170. 174354
/note="assembly-fragment"
clone_end:T7
```

BASE COUNT 46808 a 39544 c 39585 g 46362 t 2055 others  
ORIGIN

Alignment Scores:

```
Pred. No.: 4.32e+03 Length: 174354
Score: 55.00 Matches: 11
Percent Similarity: 80.00% Conservative: 1
Best Local Similarity: 73.33% Mismatches: 3
Query Match: 73.33% Indels: 0
DB: 2 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x AC108830 (1-174354)
QY 2 SerSerIleLeuCySSerProIeuProSerIeuSerIeuSerVal 16
DB 21798 AGTGAGATTCTGTGCAATGCCCTTCGCCCTTAAGTTACAAAGTC 21842
```

Search completed: July 1, 2003, 00:18:12  
Job time : 249.866 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 21:33:09 ; Search time 12.8143 Seconds  
(without alignments)  
2811.853 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75

Sequence: 1 SSSILCSPLSLSLSV 16

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=xlp  
-O=/cgn2.1/USPRO/spool/US10075846/runat\_25062003\_163647\_5141/app\_query.fasta\_1.782  
-DB=N.Geneseq\_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcio -NORML-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846 @CGN1.1.511 @runat\_25062003\_163647\_5141 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: N.Geneseq\_101002.\*  
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	90.7	39796	21 AAC61681	Nucleotide sequenc
2	53	70.7	3294	14 AAO48939	Steroid hormone re
3	53	70.7	3301	20 AAX36523	Human PPAR-delta c
4	53	70.7	3301	24 ABL65244	Lung cancer relate
5	52	69.3	90	22 ABA75659	Human foetal liver
6	52	69.3	90	22 AAK24269	Human brain expres
7	52	69.3	90	22 ABA12367	Probe #17300 for g
8	52	69.3	597	22 ABA63211	Human foetal liver
9	52	69.3	597	22 AAK1674	Human brain expres
10	52	69.3	597	22 AAI18235	Probe #8168 for ge
11	51	68.0	3599	22 AAD17468	Human transporter
12	51	68.0	20453	23 ABL03244	Drosophila melanog
13	50	66.7	418	23 ABV16725	Human prostate exp
14	50	66.7	471	23 ABV46523	Human prostate exp
15	50	66.7	1689	20 AA223024	Rat kd312 polypept
16	50	66.7	3079	20 AA223025	Rat kd312 genomic
17	49	65.3	1082	23 AAS74624	DNA encoding novel
18	49	65.3	2667	21 AAC45706	Arabidopsis thalia
19	49	65.3	3171	23 AAS74569	DNA encoding novel
20	49	65.3	3171	23 AAS75059	DNA encoding novel
21	49	65.3	112460	24 ABR83567	Human cDNA differe
22	49	65.3	249487	24 ABR85733	Mouse genomic regi
23	48	64.0	288	22 AAL16141	Human breast cance
24	48	64.0	289	22 AAL24984	Human breast cance
25	48	64.0	576	24 ABR73574	Bacillus licheniflo
26	48	64.0	668	22 AAK68950	Human immune/haema
27	48	64.0	698	22 AAI94991	Human neuroblastom
28	48	64.0	730	22 AAI95568	Human neuroblastom
29	48	64.0	827	21 AAC47631	Arabidopsis thalia
30	48	64.0	3271	19 AAV00060	Human G-protein co
31	48	64.0	4010	20 AAV72212	Rat boorta ranBC pr
32	48	64.0	5412	20 AAX21356	Human BA12 gene.
33	48	64.0	10468	22 AAS30509	DNA encoding novel
34	48	64.0	10468	22 AAS30512	DNA encoding novel
35	48	64.0	10468	22 AAL05288	Human reproductive
36	48	64.0	10468	22 AAL06291	Human reproductive
37	47.5	63.3	19408	22 AAS42003	Genomic sequence #
38	47.5	63.3	19408	22 AAK87230	Human immune/haema
39	47.5	63.3	19408	22 AAK90644	Human digestive sy
40	47	62.7	144	23 ABL26393	Drosophila melanog
41	47	62.7	255	22 AAL23844	Human breast cance
42	47	62.7	641	24 ABR82059	Novel floral meris
43	47	62.7	641	24 ABR82060	Novel floral meris
44	47	62.7	666	24 ABR82062	Novel floral meris
45	47	62.7	731	24 ABR82061	Novel floral meris

## ALIGNMENTS

RESULT 1  
AAC61681  
ID AAC61681 standard; DNA; 39796 BP.

AC AAC61681;  
19-FEB-2001 (first entry)

DE Nucleotide sequence of the human ataxia gene.

KW Human; ataxia; gene therapy; ss.

OS Homo sapiens.

FT Key Location/Qualifiers  
FT exon 29850..29921  
FT intron 29922..33025

```

FT FT      /tag- b
FT exon      33026..33155
FT          /tag- c
FT intron     33156..33444
FT          /tag- d
FT exon      33445..33514
FT          /tag- e
FT intron     33513..33751
FT          /tag- f
FT exon      33752..33975
FT          /tag- g
FT intron     33976..34114
FT          /tag- h
FT exon      34115..34195
FT          /tag- i
FT intron     34196..35759
FT          /tag- j
FT exon      35760..35901
FT          /tag- k
FT intron     35902..38781
FT          /tag- l
FT exon      38782..38996
FT          /tag- m

XX PN      WO200058461-A1.
XX PD      05-OCT-2000.
XX PF      23-MAR-2000; 2000WO-EP02600.
XX PR      26-MAR-1999; 99EP-0106343.
XX PA      (RAPP/) RAPPOLD-HOERBRAND G.
XX PI      Rappold-Hoerbrand G;
XX DR      WPI; 2000-656166/63.
XX PT      Novel nucleic acid sequence encoding human ataxia protein for screening
XX PT      compounds useful for treating disorders relating to mutations in ataxia
XX PS      gene
XX CC      Claim 6; Page 22-44; 47pp; English.
XX CC      The present sequence represents the human ataxia gene. The ataxia
XX CC      protein and polynucleotides are useful for diagnosing and treating
XX CC      disorders related to ataxia. Ataxia gene sequences are useful in
XX CC      gene therapy, and as diagnostic tools or reagents for identifying and
XX CC      characterizing genetic defect involved in the disorders and diseases
XX CC      related to ataxia.
XX SQ      Sequence 39796 BP; 10630 A; 9337 C; 9355 G; 10474 T; 0 other;

Alignment Scores:
Pred. No.:      30.5      Length:      39796
Score:          68.00     Matches:      14
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 0
Query Match:    90.67% Indels: 0
DB:            21      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AAC61681 (1-39796)
OY      2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
Db      35719 GCATCATACCTCTGCAGCCCTCTGCCATCTCTGTCTTCACTTCACTT 35763

RESULT 2
AAQ48939
ID      AAQ48939 standard; DNA; 3294 BP.
XX AC      AAQ48939;
XX XX

```

```

DT DT      13-APR-1994 (first entry)
XX XX
DE Steroid hormone receptor (NUCI) gene.
XX XX
KW Steroid; hormone; receptor; osteosarcoma; superfamily; ss.
XX XX
OS Homo sapiens.
XX XX
FH Key      Location/Qualifiers
FT CDS      338..1663
FT          /tag- a
FT          /product= Steroid hormone receptor.
FT FT
XX PN      GB2265376-A.
XX PD      29-SEP-1993.
XX PF      23-MAR-1993; 93GB-0006043.
XX PR      24-MAR-1992; 92US-0857055.
XX PA      (MERI ) MERCK & CO INC.
XX PI      Rodan GA, Rutledge SJ, Schmidt A, Vogel RL;
XX DR      WPI; 1993-305586/39.
XX DR      P-PSDB; AAR41875.
XX PT      New human steroid hormone receptor NUCI - used to assay and
XX PS      evaluate ligands binding to the receptor.
XX CC      Claim 7; Page 6-8; 62pp; English.
XX CC      The steroid hormone receptor (designated NUCI) is used in assays to
XX CC      identify and evaluate chemical entities that bind to it.
XX SQ      Sequence 3294 BP; 705 A; 1014 C; 879 G; 696 T; 0 other;

Alignment Scores:
Pred. No.:      309      Length:      3294
Score:          53.00     Matches:      12
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 1
Query Match:    70.67% Indels: 0
DB:            14      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AAQ48939 (1-3294)
OY      2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15
Db      1837 TCTTCCCTTCTCCCTTCCCTTCCCTTCCCTTCTCTCTCC 1878

RESULT 3
AAK36523
ID      AAK36523 standard; cDNA; 3301 BP.
XX AC      AAK36523;
XX DT      07-JUL-1999 (first entry)
XX DE      Human PPAR-delta coding sequence.
XX KW      Nuclear receptor agonist; antagonist; identification; PPAR;
XX KW      peroxisome proliferator activated receptor; ss.
XX OS      Homo sapiens.
XX PN      WO9918124-A1.
XX PD      15-APR-1999.
XX PF      06-OCT-1998; 98WO-US21049.
XX XX

```

XX	07-OCT-1997:	97US-0061385.	
XX			
PA	(MERI )	MERCK & CO INC.	
PI	Cummings RT, Hermes JD, Moller DE, Zhou G:		
XX			
DR	WPI: 1999-263998/22.		
DR	P-PSDB: AAY05472.		
XX			
PI	Identifying nuclear receptor agonists and antagonists		
XX			
.PS	Disclosure: Fig 10b-c: 60pp; English.		
XX			
CC	This sequence encodes the human peroxisome proliferator activated		
CC	receptor-delta (PPAR-delta).		
CC	The invention relates to a method for identifying nuclear receptor		
CC	agonists and antagonists comprises measuring fluorescent resonance energy		
CC	transfer between fluorescent-labelled nuclear receptors and		
CC	co-activators. The method can be used for identifying agonists and		
CC	antagonist of nuclear receptors.		
XX			
S0	Sequence 3301 BP: 705 A; 1014 C; 879 G; 696 T; 7 other;		
	Alignment Scores:		
	Pred. No.:	310	Length: 3301
	Score:	53.00	Matches: 12
	Percent Similarity:	92.86%	Conservative: 1
	Best local Similarity:	85.71%	Mismatches: 1
	Query Match:	70.67%	Indels: 0
	DB:	20	Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x AAX36523 (1-3301)			
OY	2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15		
	::		
Db	1837 TCTTCCTTCTCCCTTCTCCCTTCTCCCTTCTCTCTCC 1878		
RESULT 4			
ABL65244			
ID	ABL65244 standard; DNA: 3301 BP.		
XX			
AC	ABL65244:		
XX			
DT	15-MAY-2002 .(first entry)		
XX			
DE	Lung cancer related gene sequence SEQ ID NO:3581.		
XX			
KW	Human; cancer: colon; breast; ovary; oesophagus; kidney; thyroid;		
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;		
KW	gene; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	MO200194629-A2.		
XX			
PD	13-DEC-2001.		
XX			
PF	30-MAY-2001; 2001WO-US10838.		
XX			
XX	05-JUN-2000; 2000US-209473P.		
PR	05-JUN-2000; 2000US-209531P.		
PR	18-SEP-2000; 2000US-233133P.		
PR	18-SEP-2000; 2000US-233617P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	22-SEP-2000; 2000US-234567P.		
PR	22-SEP-2000; 2000US-234567P.		
PR	25-SEP-2000; 2000US-234923P.		
PR	25-SEP-2000; 2000US-234923P.		
PR	25-SEP-2000; 2000US-235077P.		
PR	25-SEP-2000; 2000US-235082P.		



XX	Homo sapiens.
OS	
XX	WO200157278-A2.
PN	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001MO-US00670.
XX	
PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.
PR	30-JUN-2000; 2000US-0608408.
PR	03-AUG-2000; 2000US-0632366.
PR	21-SEP-2000; 2000US-0234687.
PR	27-SEP-2000; 2000US-0236359.
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOLE-) MOLECULAR DYNAMICS INC.
XX	
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-488901/53.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
XX	analyzing gene expression in human cervical epithelial cells -
PS	Claim 25; SEQ ID NO 17300; 487bp; English.
XX	
CC	The present invention relates to human single exon nucleic acid probes
CC	(SENPs). The present sequence is one such probe. The SENPs are derived
CC	from human HeLa cells. The SENPs can be used to produce a single exon
CC	microarray, which can be used for measuring human gene expression in a
CC	sample derived from human cervical epithelial cells. By measuring gene
CC	expression, the probes are therefore useful in grading and/or staging
CC	of diseases of the cervix, notably cervical cancer.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 90 BP; 29 A; 16 C; 27 G; 18 T; 0 other:
XX	
Alignment Scores:	
Pred. No.:	5, 98      length: 90
Score:	52.00      Matches: 10
Percent Similarity:	92.31%      Conservative: 2
Best Local Similarity:	76.92%      Mismatches: 1
Query Match:	69.33%      Indels: 0
DB:	22      Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x AA127367 (1-90)	
OY	3 SerIleLeuCysSerProLeuPheProSerLeuSerLeuSer 15
Db	:      CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 81 TCGCATCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 43
RESULT 8	
ID	ABA63211/C
XX	ABA63211 standard; DNA: 597 BP.
AC	
DT	01-FEB-2002 (first entry)
DE	Human foetal liver single exon nucleic acid probe #11516.
KW	Human; foetal liver; gene expression; single exon nucleic acid probe; ss
OS	Homo sapiens.
PN	WO200157277-A2.
PD	09-AUG-2001.
XX	

```

PE 30-JAN-2001; 2001WO-US00669.
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 11516; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probe may be used for predicting
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;
XX
Alignment Scores:
Pred. No.: 57.2 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 22 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x ABA63211 (1-597)
Qy 3 SerTLeuGysSerProleuProSerLeuSerLeuSer 15
Db 443 TCGCTATCTTGCTCTCCTCTCCCTGCTCTCTCC 405
RESULT 9
AAK1l674/c
ID AAK1l674 standard; DNA; 597 BP.
XX
XX AAK1l674;
XX AC
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe SEQ ID NO: 11665.
XX
XX Human: brain expressed exon: gene expression analysis; probe:
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX OS Homo sapiens.
XX
XX WO200157275-A2.
XX PN 09-AUG-2001.
XX PD 30-JAN-2001; 2001WO-US00667.
XX PF
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR:
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 11516; 639bp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probe may be used for predicting
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;
XX
Alignment Scores:
Pred. No.: 57.2 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 22 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x ABA63211 (1-597)
Qy 3 SerTLeuGysSerProleuProSerLeuSerLeuSer 15
Db 443 TCGCTATCTTGCTCTCCTCTCCCTGCTCTCTCC 405

```

```
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX
PS Example 4; SEQ ID NO: 11665; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;
XX
Alignment Scores:
Pred. No.: 57.2 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x AA118235 (1-597)
OY 3 Ser1leucySSerProleuProSerleuSerleuSer 15
DB 443 TCGCTATCTGCTCTCTCTCTCTCTCTCTCTCTCC 405
RESULT 10
AA118235/c
ID AA118235 standard; DNA; 597 BP.
XX
AC AA118235;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #8168 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
```

```
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 25; SEQ ID NO 8168; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human Hela cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 597 BP; 186 A; 117 C; 116 G; 178 T; 0 other;
XX
Alignment Scores:
Pred. No.: 57.2 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x AA118235 (1-597)
OY 3 Ser1leucySSerProleuProSerleuSerleuSer 15
DB 443 TCGCTATCTGCTCTCTCTCTCTCTCTCTCTCTCC 405
RESULT 11
AAD17468/c
ID AAD17468 standard; cDNA; 3599 BP.
XX
AC AAD17468;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human transporter and ion channel-1 (TRICH-1) cDNA.
XX
KW Human; transporter and ion channel; TRICH-1; therapy; akinesia; cardiac;
KW neurological disorder; immune disorder; allergy; nootropic; dementia;
KW AIDS; Acquired Immune Deficiency Syndrome; amyotrophic lateral sclerosis;
KW cystic fibrosis; Alzheimer's disease; depression; epilepsy; scleroderma;
KW cardiomyopathy; muscular dystrophy; Crohn's disease; multiple sclerosis;
KW rheumatoid arthritis; cancer; anaemia; asthma; trauma; haemodialysis;
KW gastritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 221..1168
FT /tag= a
FT /product= "Human TRICH-1 protein"
XX
PN WO200162923-A2.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2001; 2001WO-US05942.
XX
PR 25-FEB-2000; 2000US-0184866.
PR 02-MAR-2000; 2000US-0187947.
PR 09-MAR-2000; 2000US-0188333.
PR 17-MAR-2000; 2000US-0190230.
PR 24-MAR-2000; 2000US-0192077.
PR 30-MAR-2000; 2000US-0193500.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
```

PI Yvonne H, Yang YT, Lal P, Policky JL, Nguyen DB, Au-Young J, Yao MG,  
 PI Khan FA, Melia NK, Gandhi AR, Trilbouley CM, Patterson C,  
 PI Thornton M, Greene BD, Hernandez R, Borowsky ML, Sanjanwala MS;  
 XX  
 DR WPI: 2001-582050/65.  
 DR P-PSDB: AAE10324.  
 XX  
 PT Thirteen human transporters and ion channels (referred to as TRICH-1 to  
 PT TRICH-13), useful in the diagnosis, treatment and prevention of  
 PT transport (e.g. aknesia), neurological, muscle or immunological  
 PT disorders (e.g. allergies) -  
 XX  
 PS Claim 11: Page 120-121: 131pp: English.  
 CC The present sequence is a cDNA encoding human transporters and ion  
 CC channels (TRICH-1) protein. The TRICH DNA, protein and their agonist and  
 CC antagonists are useful in the diagnosis, treatment and prevention of  
 CC transport disorders (aknesia, amyotrophic lateral sclerosis, cystic  
 CC fibrosis), neurological (Alzheimer's disease, dementia, depression,  
 CC epilepsy), muscle (e.g. cardiomyopathy, Becker's muscular dystrophy) or  
 CC immunological disorders (e.g. allergies, acquired immunodeficiency  
 CC syndrome (AIDS), Crohn's disease, rheumatoid arthritis, scleroderma,  
 CC multiple sclerosis), viral, bacterial, parasitic, protozoal and  
 CC hematologic disorders, cancer, anaemia, asthma, trauma, haemodialysis,  
 CC gastritis and inflammation.  
 XX  
 SQ Sequence 3599 BP; 1041 A; 705 C; 775 G; 1077 T; 1 other:  
 Alignment Scores:  
 Pred. No.: 489 Length: 3599  
 Score: 52.00 Matches: 8  
 Percent Similarity: 92.86% Conservative: 5  
 Best Local Similarity: 57.14% Mismatches: 1  
 Query Match: 69.33% Indels: 0  
 DB: 22 Gaps: 0  
 US-10-075-846-4\_COPY192\_207 (1-16) x AAD17468 (1-3599)  
 QY 1 SerSerSerIleLeucySerProLeuProSerLeuSerLeu 14  
 Db 3459 GCCAAGCTCGTTTGTGCTCACCCAGCCCTCTATCCACTT 3418  
 RESULT 12  
 ABL03244/c  
 ID ABL03244 standard; cDNA: 20453 BP.  
 AC  
 XX ABL03244:  
 XX  
 DT 26-MAR-2002 (first entry)  
 DE XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 4214.  
 XX  
 KW Drosophila: developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 OS  
 XX  
 MO200171042-A2.  
 PD  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US09231.  
 XX  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 XX  
 XX Venter JC, Adams M, Li PWD, Myers EW,  
 DR WPI: 2001-656860/75.  
 DR P-PSDB: ABB59141.  
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
PS Claim 1; SEQ ID NO 4214; 21np + Sequence Listing; English.  
XX

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC disclosable, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL01840-ABL01675), expressed DNA  
CC sequences (ABL01840-ABL01675) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SO Sequence 20453 BP; 6125 A; 4190 C; 4122 G; 6016 T; 0 other;

Alignment Scores:  
Pred. No.: 5.53e+03 Length: 20453  
Score: 51.00 Matches: 10  
Percent Similarity: 80.00% Conservative: 2  
Best Local Similarity: 66.67% Mismatches: 3  
Query Match: 68.00% Indels: 0  
Gaps: 0  
23

US-10-075-846-4\_COPY\_192\_207 (1-16) x ABL03244 (1-20453)

OY 1 SerSerSertleucCysSerPioleuProSerleuSerleuSer 15  
Db 18769 TCGTCCTCTCTTGGCTCTCCGCCGCCTGTGTCGCT 18725  
|||||||::: ||||| ||||| ||:::

RESULT 13  
ABV16725/C  
ID ABV16725 standard; cDNA; 418 BP.  
XX  
AC ABV16725;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 16716.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Schlegel R, Endege WO, Monahan JF;  
DR WPI; 2001-662795/76.  
XX

Novel isolated nucleic acid molecule associated with cancerous state of  
prostate cells and correlating with presence of prostate cancer, useful  
for detecting presence of prostate cancer, stage of prostate cancer -  
Claim 1; Page 2785; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 418 BP; 133 A; 71 C; 103 G; 111 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 75.7 Length: 418  
Score: 50.00 Matches: 9  
Percent Similarity: 92.86% Conservative: 4  
Best Local Similarity: 64.29% Mismatches: 1  
Query Match: 66.67% Indels: 0  
DB: 23 Gaps: 0  
  
US-10-075-846-4\_COPY\_192\_207 (1-16) x ABV16725 (1-418)  
QY 1 SerSerSerIleLeuCySerProleuProSerIleuSerIleu 14  
DB 161 TCAGCACCTGTCTGTCTCCCTCCCTACTGTACGCTT 120  
  
RESULT 14  
ID ABV46523/C  
AC ABV46523 standard; cDNA: 471 BP.  
XX  
AC ABV46523;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker CDNA 46514.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189682P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer  
XX  
PS Claim 1; Page 9177; 11750pp; English.  
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 471 BP; 149 A; 89 C; 115 G; 118 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 87.3 Length: 471  
Score: 50.00 Matches: 9  
Percent Similarity: 92.86% Conservative: 4  
Best Local Similarity: 64.29% Mismatches: 1  
Query Match: 66.67% Indels: 0  
DB: 23 Gaps: 0  
  
US-10-075-846-4\_COPY\_192\_207 (1-16) x ABV46523 (1-471)  
QY 1 SerSerSerIleLeuCySerProleuProSerIleuSerIleu 14  
DB 214 TCAGCACCTGTCTGTCTCCCTCCCTACTGTACGCTT 173  
  
RESULT 15  
ID AA223024  
AC AA223024 standard; cDNA: 1689 BP.  
XX  
AC AA223024;  
XX  
DT 17-JAN-2000 (first entry)  
XX  
DE Rat kd312 polypeptide encoding cDNA.  
XX  
KW kd312 polypeptide; cancer; immunodeficiency disorder; AIDS; stroke;  
KW heart attack; head trauma; neurodegenerative disease; rat;  
KW Parkinson's disease; Alzheimer's disease; ss.  
XX  
OS Rattus sp.  
XX  
PN WO9950288-A2.  
XX  
PD 07-OCT-1999.  
XX  
PF 30-MAR-1999; 99WO-US06993.  
XX  
PR 31-MAR-1998; 98US-0053374.  
XX  
PA (AMGE-) AMGEN INC.  
XX  
PI Yen K;  
XX  
PT WPI; 1999-601322/51.  
XX  
DR P-PSDB; AAY42694.  
XX  
PT kd312 polypeptides useful for treating diseases and disorders  
XX associated with alterations in cell proliferation and cell death  
XX  
PS Claim 2; Fig 9; 85pp; English.  
XX  
PT The invention provides nucleic acid molecules encoding human and rat  
XX kd312 polypeptides. The kd312 polypeptides can be expressed by standard  
XX recombinant methodology. The kd312 sequences, and the antibodies against  
XX the proteins may be used to treat or diagnose the presence or progression  
XX of various diseases, e.g. cancer, immunodeficiency disorders (e.g. AIDS),



CC stroke, heart attack, head trauma, and neurodegenerative diseases (e.g.  
CC Parkinson's disease and Alzheimer's disease). The present sequence  
CC represents the rat kd312 cDNA sequence.

XX  
SQ Sequence 1689 BP; 417 A; 462 C; 434 G; 376 T; 0 other;

## Alignment Scores:

Pred. No.:	401	Length:	1689
Score:	50.00	Matches:	10
Percent Similarity:	83.33%	Conservative:	0
Best Local Similarity:	83.33%	Mismatches:	2
Query Match:	66.67%	Indels:	0
DB:	20	Gaps:	0

US-10-075-846-4\_COPY\_192\_207 (1-16) x AA223024 (1-1689)

OY 2 SerSerIleIeuCysSerProIeuProSerIeuSer 13  
||| |||||  
DB 66 TCCGGCGCCCTCTGCACGCCCTCTACCTCTCTCAGC 101

Search completed: June 30, 2003, 23:13:47  
Job time : 18.8143 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:02:29 ; Search time 3.40045 Seconds

(without alignments)  
1442.995 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75

Sequence: 1 SSILCSPLSLSLV 16

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2_1/USPTO_spool/US10075846/runat_25062003.163649.5167/app_query.fasta_1.782  
-DB=Issued_Patents_NA -OEXT=fastcap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10075846.ecgn.1.1.84.ernat.25062003.163649.5167 -MCPN=6 -ICU=3  
-NO_MMAP -LARGEMEMORY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

## Database :

- Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	66.7	1689	4	US-09-053-374A-4
2	50	66.7	3079	4	US-09-053-374A-6
3	48	64.0	3271	2	US-08-852-806-1
4	48	64.0	3271	3	US-09-163-668-1
5	46	61.3	724	4	US-09-535-008-40
6	46	61.3	1331	4	US-09-179-221D-3
7	46	61.3	3191	4	US-09-453-702B-70
8	46	61.3	4680	1	US-08-254-358-1
9	46	61.3	4680	1	US-08-475-391-1
10	46	61.3	4680	2	US-08-709-609-1
11	46	61.3	4680	5	PCT-US95-07178-1
12	46	61.3	4910	2	US-08-331-384-2

13	46	61.3	4910	2	US-08-836-087-2	Sequence 2, Appl1
14	46	61.3	4910	4	US-09-246-320-2	Sequence 2, Appl1
15	46	61.3	4910	4	US-09-546-738-2	Sequence 2, Appl1
16	46	61.3	7557	4	US-09-770-315-3	Sequence 3, Appl1
17	46	61.3	8698	4	US-09-770-315-2	Sequence 2, Appl1
18	46	61.3	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
19	46	61.3	4411529	4	US-09-103-840A-1	Sequence 1, Appl1
20	45	60.0	363	4	US-08-580-031A-13	Sequence 13, Appl1
21	45	60.0	1876	4	US-08-840-767-5	Sequence 5, Appl1
22	45	60.0	2234	4	US-09-087-134-7	Sequence 7, Appl1
23	45	60.0	2303	3	US-09-289-376-1	Sequence 1, Appl1
24	45	60.0	2333	2	US-08-732-028-1	Sequence 1, Appl1
25	45	60.0	2333	4	US-09-183-228-1	Sequence 1, Appl1
26	45	60.0	2976	4	US-09-221-017B-16	Sequence 36, Appl1
27	45	60.0	50000	4	US-09-146-053-3	Sequence 3, Appl1
28	44	58.7	517	4	US-09-288-143-43	Sequence 43, Appl1
29	44	58.7	1140	1	US-08-289-653-2	Sequence 2, Appl1
30	44	58.7	2845	1	US-08-289-653-1	Sequence 1, Appl1
31	44	58.7	4609	4	US-09-322-478-21	Sequence 21, Appl1
32	44	58.7	9829	4	US-09-322-478-19	Sequence 19, Appl1
33	44	58.7	11707	4	US-09-136-574A-1	Sequence 1, Appl1
34	44	58.7	12286	4	US-09-322-478-17	Sequence 17, Appl1
35	44	58.7	152331	3	US-09-128-155-16	Sequence 16, Appl1
36	44	58.7	176373	3	US-09-128-155-17	Sequence 17, Appl1
37	43.5	58.0	1906	4	US-09-031-962D-5	Sequence 5, Appl1
38	43.5	58.0	3153	4	US-09-175-928-9	Sequence 9, Appl1
39	43	57.3	2820	5	PCT-US93-11725-1	Sequence 1, Appl1
40	43	57.3	3074	5	PCT-US93-11725-3	Sequence 3, Appl1
41	43	57.3	7015	4	US-09-177-249-6	Sequence 6, Appl1
42	43	57.3	29629	4	US-09-729-995-3	Sequence 3, Appl1
43	42	56.0	185	4	US-09-402-668-11	Sequence 11, Appl1
44	42	56.0	1109	1	US-08-183-214-11	Sequence 11, Appl1
45	42	56.0	1168	4	US-09-484-970B-89	Sequence 89, Appl1

## ALIGNMENTS

```
RESULT 1  
US-09-053-374A-4  
; Sequence 4, Application US/09053374A  
; Patent No. 6462177  
; GENERAL INFORMATION:  
; APPLICANT: YEN, KWANG-MU  
; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: AMGEN INC.  
; STREET: ONE AMGEN CENTER DRIVE  
; CITY: THOUSAND OAKS  
; STATE: CA  
; COUNTRY: US  
; ZIP: 91320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/053,374A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: COOK, ROBERT R.  
; REGISTRATION NUMBER: 31,602  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1689 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:
```

```

: NAME/KEY: CDS
: LOCATION: 132..971
US-09-053-374A-4
Alignment Scores:
Pred. NO.: 71.8 Length: 1689
Score: 50.00 Matches: 10
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 2
Query Match: 66.67% Indels: 0
DB: 4 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x US-09-053-374A-4 (1-1689)
OY 2 SerSerIleLeuGysSerProLeuProSerLeuSer 13
Db 66 TCCGGCGCCCTCTGCAGCCCTCTACTCTCTCTCAGC 101
RESULT 2
US-09-053-374A-6
: Sequence 6, Application US/09053374A
: Patent No. 6462177
: GENERAL INFORMATION:
: APPLICANT: YEN, KWANG-MU
: TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: AMGEN INC.
: STREET: ONE AMGEN CENTER DRIVE
: CITY: THOUSAND OAKS
: STATE: CA
: COUNTRY: US
: ZIP: 91320
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/053,374A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: COOK, ROBERT R.
: REGISTRATION NUMBER: 31,602
: REFERENCE/DOCKET NUMBER: A-514
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3079 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-09-053-374A-6
Alignment Scores:
Pred. NO.: 147 Length: 3079
Score: 50.00 Matches: 10
Percent Similarity: 83.33% Conservative: 0
Best Local Similarity: 83.33% Mismatches: 2
Query Match: 66.67% Indels: 0
DB: 4 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x US-09-053-374A-6 (1-3079)
OY 2 SerSerIleLeuGysSerProLeuProSerLeuSer 13
Db 237 TCCGGCGCCCTCTGCAGCCCTCTACTCTCTCTCAGC 272
RESULT 3
US-08-852-806-1
: Sequence 1, Application US/08852806
: Patent No. 5874245
: GENERAL INFORMATION:
: APPLICANT: Shoji Fukusumi
: TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: U.S.A.
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ Version 1.5
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/852,806
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/017,915
: FILING DATE: 16 MAY 1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Han, William T
: REGISTRATION NUMBER: 34,344
: REFERENCE/DOCKET NUMBER: TAK50002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5219
: TELEFAX: 610-270-5090
: TELEX:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3271 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE:
: ORIGINAL SOURCE:
US-08-852-806-1
Alignment Scores:
Pred. NO.: 321 Length: 3271
Score: 48.00 Matches: 9
Percent Similarity: 90.91% Conservative: 1
Best Local Similarity: 81.82% Mismatches: 1
Query Match: 64.00% Indels: 0
DB: 2 Gaps: 0
US-10-075-846-4_COPY_192_207 (1-16) x US-08-852-806-1 (1-3271)
OY 5 LeuGysSerProLeuProSerLeuSerLeuSer 15
Db 1064 CTGTGGCCGCCCTCTACCCAGCCTCTCAGCTGAGC 1096
RESULT 4
US-09-163-669-1
: Sequence 1, Application US/09163669
: Patent No. 611076
: GENERAL INFORMATION:
: APPLICANT: FUKUSUMI, SHOJI
: APPLICANT: HINUMA, SHUJI
: APPLICANT: FUJII, RYO
: TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED
: TITLE OF INVENTION: RECEPTOR (HIBCD07)
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ratner & Prestlia
: STREET: P.O. Box 980

```

CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/163,669  
FILING DATE: 30-SEP-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/852,806  
FILING DATE: 07-MAY-1997  
APPLICATION NUMBER: 60/017,915  
FILING DATE: 16-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: TAK-50002-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3271 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-09-163-669-1

Alignment Scores:  
Pred. No.: 321 Length: 3271  
Score: 48.00 Matches: 9  
Percent Similarity: 90.91% Conservative: 1  
Best Local Similarity: 81.82% Mismatches: 1  
Query Match: 64.00% Indels: 0  
DB: 3 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-163-669-1 (1-3271)  
QY 5 LeucySerProLeuProSerLeuSerLeuSer 15  
DB 1064 CTGTGCGCCGCCCTACCCAGCCTCCAGCTGAGC 1096

RESULT 5  
US-09-535-008-40  
Sequence 40, Application US/09535008  
Patent No. 6465629  
GENERAL INFORMATION:  
APPLICANT: Wong, Alexander K.C.  
APPLICANT: Tavligian, Sean V.  
APPLICANT: Teng, David H.F.  
TITLE OF INVENTION: BRG1 IS A TUMOR SUPPRESSOR THAT IS MUTATED IN PROSTATE  
TITLE OF INVENTION: AND OTHER CANCER TYPES  
FILE REFERENCE: 2318-259  
CURRENT APPLICATION NUMBER: US/09/535,008  
CURRENT FILING DATE: 2000-03-23  
EARLIER APPLICATION NUMBER: U.S. 60/125,806  
EARLIER FILING DATE: 1999-03-23  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: Patencin Ver. 2.0  
SEQ ID NO 40  
LENGTH: 724  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-535-008-40

Alignment Scores:

Pred. No.: 109 Length: 724  
Score: 46.00 Matches: 8  
Percent Similarity: 84.62% Conservative: 3  
Best Local Similarity: 61.34% Mismatches: 2  
Query Match: 61.33% Indels: 0  
DB: 4 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-535-008-40 (1-724)  
QY 4 lleucySerProLeuProSerLeuSerLeuSerVal 16  
DB 681 GTTCTGTGCTCCCGCCGCTCCCTTCACCTACCTC 719

RESULT 6  
US-09-179-221D-3/C  
Sequence 3, Application US/09179221D  
Patent No. 6291168  
GENERAL INFORMATION:  
APPLICANT: Musso, Richard  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES DIAGNOSTIC FOR PATHOGENIC E.  
TITLE OF INVENTION: COLI 0157, METHODS OF IDENTIFICATION AND KIT THEREFOR  
FILE REFERENCE: 922,6496P  
CURRENT APPLICATION NUMBER: US/09/179,221D  
CURRENT FILING DATE: 1998-10-27  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: No. 6291168epad  
SEQ ID NO 3  
LENGTH: 1331  
TYPE: DNA  
ORGANISM: E. coli  
US-09-179-221D-3

Alignment Scores:  
Pred. No.: 225 Length: 1331  
Score: 46.00 Matches: 8  
Percent Similarity: 80.00% Conservative: 4  
Best Local Similarity: 53.33% Mismatches: 3  
Query Match: 61.33% Indels: 0  
DB: 4 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-179-221D-3 (1-1331)  
QY 2 serSerlleucySerProLeuProSerLeuSerLeuSerVal 16  
DB 102 ACATCAATCATCTGTTCTCCATGACTAGTCTAAACTAGATTT 58

RESULT 7  
US-09-453-702B-70  
Sequence 70, Application US/09453702B  
Patent No. 6365723  
GENERAL INFORMATION:  
APPLICANT: Blattner, Frederick R.  
APPLICANT: Perna, Nicole T.  
APPLICANT: Plunkett, Guy  
TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESS: Quarles & Brady  
STREET: 1 South Plunkney Street  
CITY: Madison  
STATE: WI  
COUNTRY: US  
ZIP: 53701-2113  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 8.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/453,702B  
FILING DATE: 03-Dec-1999

```
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 3191
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-453-702B-70

Alignment Scores:
Pred. No.: 635          Length: 3191
Score: 46.00          Matches: 8
Percent Similarity: 80.008      Conservative: 4
Best Local Similarity: 53.338    Mismatches: 3
Query Match: 61.338             Indels: 0
DB: 4                          Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-453-702B-70 (1-3191)
QY 2 SerSerIleleuCySerProleuProSerleuSerleuSerVal 16
Db 547 ACATCAATCATCTGCTCCAACTAGCTAATAAAGTAGTATT 591

RESULT 8
US-08-254-358-1
Sequence 1, Application US/08254358
Patent No. 5658785
GENERAL INFORMATION:
APPLICANT: Johnson, Phillip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/254,358
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5658785and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
```

```
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-254-358-1

Alignment Scores:
Pred. No.: 1e+03          Length: 4680
Score: 46.00          Matches: 10
Percent Similarity: 73.338      Conservative: 1
Best Local Similarity: 66.674    Mismatches: 4
Query Match: 61.338             Indels: 0
DB: 1                          Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-08-254-358-1 (1-4680)
QY 1 SerSerIleleuCySerProleuProSerleuSerleuSer 15
Db 775 ACATCCCAATTAATCTGCTCCCAAAACCACTAGCTCCAGT 819

RESULT 9
US-08-475-391-1
Sequence 1, Application US/08475391
Patent No. 5786211
GENERAL INFORMATION:
APPLICANT: Johnson, Phillip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,391
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/254,358
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5786211and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-475-391-1

Alignment Scores:
Pred. No.: 1e+03          Length: 4680
Score: 46.00          Matches: 10
Percent Similarity: 73.338      Conservative: 1
Best Local Similarity: 66.674    Mismatches: 4
Query Match: 61.338             Indels: 0
DB: 1                          Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-08-475-391-1 (1-4680)
```

QY 1 SerSerSerIleuCySerProleuProSerIleuSer 15  
Db 775 ACATCCCAATTACTTGCTCCCAAAACCAACGAGCTGCTCCAGT 819

RESULT 10  
US-08-709-609-1  
Sequence 1, Application US/08709609  
Patent No. 5858775  
GENERAL INFORMATION:  
APPLICANT: Johnson, Phillip R.  
TITLE OF INVENTION: Adeno-Associated Virus Materials and  
METHODS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois  
ZIP: 60606  
COUNTRY: USA

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,609  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: No. 5858775and, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31975  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

US-08-709-609-1

Alignment Scores:  
Pred. No.: 1e+03 Length: 4680  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-08-709-609-1 (1-4680)

QY 1 SerSerSerIleuCySerProleuProSerIleuSer 15  
Db 775 ACATCCCAATTACTTGCTCCCAAAACCAACGAGCTGCTCCAGT 819

RESULT 11  
PCT-US95-07178-1  
Sequence 1, Application PC/TUS9507178  
GENERAL INFORMATION:  
APPLICANT: Johnson, Phillip R.  
TITLE OF INVENTION: Adeno-Associated Virus Materials and  
METHODS  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 S. Wacker Drive  
CITY: Chicago  
STATE: Illinois

COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07178  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Noland, Greta E.  
REGISTRATION NUMBER: 35,302  
REFERENCE/DOCKET NUMBER: 31975  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 474-6300  
TELEFAX: (312) 474-0448  
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4680 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

PCT-US95-07178-1

Alignment Scores:  
Pred. No.: 1e+03 Length: 4680  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 5 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x PCT-US95-07178-1 (1-4680)

QY 1 SerSerSerIleuCySerProleuProSerIleuSer 15  
Db 775 ACATCCCAATTACTTGCTCCCAAAACCAACGAGCTGCTCCAGT 819

RESULT 12  
US-08-331-384-2/C  
Sequence 2, Application US/08331384  
Patent No. 5856152  
GENERAL INFORMATION:  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and Methods  
OF USE THEREOF  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331,384  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: UPNG1149USA

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-08-331-384-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-08-331-384-2 (1-4910)  
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSer 15  
Db 2062 ACATCCCAATTACTGCTCCCAAAACCGCTGAGCTCCAGT 2018

RESULT 13  
US-08-836-087-2/c  
Sequence 2, Application US/08836087  
Patent No. 5871982  
GENERAL INFORMATION:  
APPLICANT: Trustees of University of Pennsylvania  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/836.087  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/331.384  
FILING DATE: 28-OCT-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: GNVPN.007PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-08-836-087-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 2 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-08-836-087-2 (1-4910)  
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSer 15  
Db 2062 ACATCCCAATTACTGCTCCCAAAACCGCTGAGCTCCAGT 2018

RESULT 14  
US-09-246-320-2/c  
Sequence 2, Application US/09246320  
Patent No. 6251677  
GENERAL INFORMATION:  
APPLICANT: Trustees of University of Pennsylvania  
APPLICANT: Wilson, James M.  
APPLICANT: Kelley, William M.  
APPLICANT: Fisher, Krishna J.  
TITLE OF INVENTION: Hybrid Adenovirus-AAV Vector and  
TITLE OF INVENTION: Methods of Use Thereof  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/246.320  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/836.087  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: GNVPN.007PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4910 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
US-09-246-320-2

Alignment Scores:  
Pred. No.: 1.06e+03 Length: 4910  
Score: 46.00 Matches: 10  
Percent Similarity: 73.33% Conservative: 1  
Best Local Similarity: 66.67% Mismatches: 4  
Query Match: 61.33% Indels: 0  
DB: 4 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-246-320-2 (1-4910)  
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSer 15  
Db 2062 ACATCCCAATTACTGCTCCCAAAACCGCTGAGCTCCAGT 2018





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:07:05 ; Search time 8.51902 Seconds  
(without alignments)  
2788.018 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSIICSPPLSLSLV 16

Scoring table: BIOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1055720 segs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-O/cgn2\_1/USFPO.spool/US10075846/runat\_25062003\_163649\_5181/app.query.fasta.1.762  
-DB=Published.Applications\_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62  
-TRANS=Numan40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US10075846 @CGN.1.1.102 @runat\_25062003\_163649\_5181  
-NCPU=6 -ICPU=3 -NO\_MAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOC -DEV=TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published.Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCR\_NEM\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEM\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEM\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCRUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEM\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEM\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEM\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEM\_PUB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	1640	9	US-10-075-846-3
2	53	70.7	2982	9	US-08-938-842A-93
3	53	70.7	3301	10	US-09-954-456-554
4	52	69.3	90	10	US-09-864-761-30015

C	5	52	69.3	597	10	US-09-864-761-13469	Sequence 13469, A
	6	49	65.3	249487	9	US-10-026-188-3	Sequence 3, Appl1
	7	48	64.0	497	10	US-09-783-590-11273	Sequence 11273, A
C	8	48	64.0	576	10	US-09-974-300-865	Sequence 865, App
C	9	48	64.0	720	9	US-10-138-846-12894	Sequence 12894, A
C	10	48	64.0	791	9	US-10-198-846-7241	Sequence 7241, App
	11	48	64.0	5399	9	US-10-225-567A-343	Sequence 343, App
	12	48	64.0	10468	9	US-09-764-891-8976	Sequence 8976, App
	13	48	64.0	10468	9	US-09-764-891-8976	Sequence 8976, App
	14	47	62.7	263	9	US-09-918-995-8979	Sequence 8979, App
C	15	47	62.7	381	10	US-09-983-965-3940	Sequence 3940, App
	16	47	62.7	771	9	US-09-764-891-725	Sequence 725, App
C	17	47	62.7	4000	10	US-09-880-107-3430	Sequence 3430, App
	18	47	62.7	110079	9	US-10-175-523-96	Sequence 96, Appl
C	19	47	62.7	174424	10	US-09-967-768A-314	Sequence 314, App
C	20	47	62.7	465237	10	US-09-933-267A-1	Sequence 1, Appl1
C	21	46	61.3	231	10	US-09-864-761-17984	Sequence 17984, A
C	22	46	61.3	254	9	US-09-728-444-304	Sequence 304, App
C	23	46	61.3	289	9	US-09-991-936-556	Sequence 556, App
	24	46	61.3	408	9	US-09-918-995-23458	Sequence 23458, A
	25	46	61.3	424	10	US-09-783-590-2728	Sequence 2728, App
	26	46	61.3	479	10	US-09-864-761-1221	Sequence 1221, App
C	27	46	61.3	1363	10	US-09-880-192-27	Sequence 27, Appl
C	28	46	61.3	1591	10	US-09-925-300-290	Sequence 290, App
	29	46	61.3	1611	9	US-09-792-630-16	Sequence 16, Appl
	30	46	61.3	1611	9	US-09-792-630-16	Sequence 16, Appl
	31	46	61.3	1611	9	US-10-080-376-16	Sequence 16, Appl
	32	46	61.3	1611	9	US-10-080-376-16	Sequence 16, Appl
	33	46	61.3	1611	9	US-09-953-351-16	Sequence 16, Appl
	34	46	61.3	1611	9	US-09-953-351-16	Sequence 16, Appl
	35	46	61.3	1611	9	US-10-082-671-22	Sequence 22, Appl
	36	46	61.3	1611	9	US-10-082-671-22	Sequence 22, Appl
	37	46	61.3	1611	9	US-10-097-100-16	Sequence 16, Appl
	38	46	61.3	1611	9	US-10-097-100-16	Sequence 16, Appl
	39	46	61.3	1866	9	US-09-792-630-12	Sequence 12, Appl
	40	46	61.3	1866	9	US-09-792-630-12	Sequence 12, Appl
	41	46	61.3	1866	9	US-10-080-376-2	Sequence 2, Appl1
	42	46	61.3	1866	9	US-10-080-376-2	Sequence 2, Appl1
	43	46	61.3	1866	9	US-09-953-351-2	Sequence 2, Appl1
	44	46	61.3	1866	9	US-09-953-351-2	Sequence 2, Appl1
	45	46	61.3	1866	9	US-10-082-671-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-10-075-846-3  
: Sequence 3, Application US/10075846  
: Publication No. US20030032608A1  
: GENERAL INFORMATION:  
: APPLICANT: Bristol-Myers Squibb Company  
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
: FILE REFERENCE: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THERE  
: CURRENT APPLICATION NUMBER: US/10/075, 846  
: PRIOR FILING DATE: 2002-02-13  
: PRIOR APPLICATION NUMBER: US 60/269, 535  
: NUMBER OF SEQ ID NOS: 81  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO 3  
: LENGTH: 1640  
: TYPE: DNA  
: ORGANISM: homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (1)..(1293)  
US-10-075-846-3

Alignment Scores:  
Pred. No.: 0.00593 Length: 1640  
Score: 75.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-10-075-846-3 (1-1640)

OY 1 SerSerSerIleLeuGysSerProLeuProSerLeuSerLeuSerVal 16  
DB 574 AGCTCATCATCATCTGACGCCCTCTCCATCTCTGTCATCTTCAGTT 621

## RESULT 2

US-09-938-842A-93/C

Sequence 93, Application US/09938842A  
Patent No. US20020160378A1

GENERAL INFORMATION:

APPLICANT: Harper, Jeff

APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

FILE REFERENCE: SAME, AND METHODS OF USE

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO 93

LENGTH: 2982

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-93

Alignment Scores:

Pred. No.: 40.1 Length: 2982

Score: 53.00 Matches: 10

Percent Similarity: 81.25% Conservative: 3

Best Local Similarity: 62.50% Mismatches: 3

Query Match: 70.67% Indels: 0

DB: 9 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-938-842A-93 (1-2982)

OY 1 SerSerSerIleLeuGysSerProLeuProSerLeuSerLeuSerVal 16  
DB 2893 TCTTATTGAGTAGCTGTGCACCTCTCCCTCAATCTGTATCAGTT 2846

## RESULT 3

US-09-954-456-554  
Sequence 554, Application US/09954456  
Patent No. US20020115057A1

GENERAL INFORMATION:

APPLICANT: Young, Paul

TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

FILE REFERENCE: 689290-76

CURRENT FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/09/954,456

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/233,617

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/234,052

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,923

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/60/235,134

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/60/235,637

PRIOR FILING DATE: 2000-09-26

PRIOR APPLICATION NUMBER: US/60/235,638

PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,840  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,863  
PRIOR FILING DATE: 2000-09-27  
NUMBER OF SEQ ID NOS: 2276  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 554  
LENGTH: 3301  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: n-a,t,g or c  
US-09-954-456-554

## Alignment Scores:

Pred. No.: 44.8 Length: 3301

Score: 53.00 Matches: 12

Percent Similarity: 92.86% Conservative: 1

Best Local Similarity: 85.71% Mismatches: 1

Query Match: 70.67% Indels: 0

DB: 10 Gaps: 0

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-954-456-554 (1-3301)

OY 2 SerSerSerIleLeuGysSerProLeuProSerLeuSerLeuSer 15  
DB 1837 TCTTCCCTTCTCTCCCTTCTCCCTTCTCTCTCTCC 1878

## RESULT 4

US-09-864-761-30015/C

Sequence 30015, Application US/09864761  
Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

FILE REFERENCE: Aegm1ca-X-1

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 30015
: LENGTH: 90
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL138761.3
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
: OTHER INFORMATION: NT HIT: AB021490.2, EVALU6 7.40e-01
: OTHER INFORMATION: EST_HUMAN HIT: BE612820.1, EVALU6 1.10e+00
US-09-864-761-30015
```

```

Alignment Scores:
Pred. No.: 1.25 Length: 90
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 10 Gaps: 0
```

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-864-761-30015 (1-90)

```

QY 3 SerileuCySerProleuProSerleuSerleuSer 15
||||| ||||||||| ||||||||| |||||||||
DB 81 TCGCATCTTGCTGCTCTCTCTCTCTCTCTCTCTCC 43
```

```

RESULT 5
US-09-864-761-13469/C
: Sequence 13469, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aecmics-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
```

```

: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 13469
: LENGTH: 597
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL138761.3
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.6
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
US-09-864-761-13469
```

```

Alignment Scores:
Pred. No.: 9.97 Length: 597
Score: 52.00 Matches: 10
Percent Similarity: 92.31% Conservative: 2
Best Local Similarity: 76.92% Mismatches: 1
Query Match: 69.33% Indels: 0
DB: 10 Gaps: 0
```

US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-864-761-13469 (1-597)

```

QY 3 SerileuCySerProleuProSerleuSerleuSer 15
||||| ||||||||| ||||||||| |||||||||
DB 443 TCGCATCTTGCTGCTCTCTCTCTCTCTCTCTCTCC 405
```

```

RESULT 6
US-10-026-188-3
: Sequence 3, Application US/10026188
: Patent No. US20020164645A1
: GENERAL INFORMATION:
: APPLICANT: Zuker, Charles S.
: APPLICANT: Zhang, Yifeng
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
: FILE REFERENCE: 02307E-114910US
: CURRENT APPLICATION NUMBER: US/10/026,188
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: US 60/259,379
: PRIOR FILING DATE: 2000-12-29
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 249487
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: OTHER INFORMATION: mouse genomic region containing Itprcs
US-10-026-188-3

Alignment Scores:
Pred. No.: 2.25e+04 Length: 249487
```

```

Percent Similarity:      49.00%    Matches:      11
Score:                   80.00%    Conservative: 1
Best local similarity:   73.33%    Mismatches:  3
Query Match:            65.33%     Indels:       0
DB:                      9          Gaps:         0

US-10-075-846-4_COPY_192_207 (1-16) x US-10-026-188-3 (1-249487)
Oy      1 SerSerIleLeuYsSerProIeuProSerIeuSerIauSer 15
Db      191671 TCCTCCTCCTCCTCCTCCTCCTCCCTCCCTCCCTCCTCCTCCTCC 191715

RESULT 7
US-09-783-590-11273
Sequence 11273, Application US/09783590
Patent No. US20020110850A1
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Haseltine, William A.
APPLICANT: Li, Haodong
APPLICANT: Rosen, Craig A.
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
FILE REFERENCE: PO-16.2c1
CURRENT APPLICATION NUMBER: US/09/783,590
PRIOR FILING DATE: 2000-02-15
PRIOR APPLICATION NUMBER: 08/420,856
PRIOR FILING DATE: 1995-04-12
PRIOR APPLICATION NUMBER: 08/346,731
PRIORITY FILING DATE: 1994-11-21
NUMBER OF SEQ ID NOS: 12485
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 11273
LENGTH: 497
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (54)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (77)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (121)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (288)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (305)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (317)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (389)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (397)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (408)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (422)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (430)
OTHER INFORMATION: n equals a,t,g, or c

```

```

NAME/KEY: misc feature
LOCATION: (432)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (436)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (447)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (448)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (453)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (460)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (468)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (472)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (474)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (477)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (481)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (483)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (486)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (487)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (490)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (497)
OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11273

Alignment Scores:
Pred. No.: 36
Score: 48.00
Percent Similarity: 64.62%
Percent Local Similarity: 64.23%
Query Match: 64.00%
DB: 10
Length: 497
Matches: 9
Conservative: 2
Mismatches: 2
Indels: 0
Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-783-590-11273 (1-497)
Oy 2 SerSerlleleuCySsSerProleuProSerleuSerleu 14
Db 164 TCTATTCTTTGTGTCGCCACGTGCTCCCTCACTCCCAATT 202

RESULT 8
US-09-974-300-865/C
Sequence 865 Application US/09974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Berka, Randy M.
APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
TITLE OF INVENTION: Expression

```

```

; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974.300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680.598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279.526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 865
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-865

Alignment Scores:
Pred. No.: 42.3      Length: 576
Score: 48.00      Matches: 9
Percent Similarity: 85.71%      Conservative: 3
Best Local Similarity: 64.29%      Mismatches: 2
Query Match: 64.00%      Indels: 0
DB: 10      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-09-974-300-865 (1-576)
Oy 2 SerSerIleuCySerProleuProSerIleuSer 15
DB 557 CCTTCTACTATATCTGCTCCGCGCTCTCTTTCAATCGCG 516

RESULT 9
US-10-198-846-12894/C
; Sequence 12894, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198.846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306.220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12894
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; LOCATION: 1, 2, 3, 716, 717, 718, 719, 720
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-12894

Alignment Scores:
Pred. No.: 54      Length: 720
Score: 48.00      Matches: 8
Percent Similarity: 80.00%      Conservative: 4
Best Local Similarity: 53.33%      Mismatches: 3
Query Match: 64.00%      Indels: 0
DB: 9      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-10-198-846-12894 (1-720)
Oy 1 SerSerIleuCySerProleuProSerIleuSer 15
DB 316 ACTACGGGTCTATCTGCTCGCGGCCCATCAGCTCCATGCGC 272

RESULT 10

US-10-198-846-7241/C
; Sequence 7241, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198.846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306.220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 7241
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc-feature
; LOCATION: 282, 300, 554, 592, 650, 707, 720, 728, 755, 767, 775, 791
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7241

Alignment Scores:
Pred. No.: 59.9      Length: 791
Score: 48.00      Matches: 8
Percent Similarity: 80.00%      Conservative: 4
Best Local Similarity: 53.33%      Mismatches: 3
Query Match: 64.00%      Indels: 0
DB: 9      Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x US-10-198-846-7241 (1-791)
Oy 1 SerSerIleuCySerProleuProSerIleuSer 15
DB 81 ACTACGGGTCTATCTGCTCGCGGCCCATCAGCTCCATGACG 37

RESULT 11
US-10-225-567A-343
; Sequence 343, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPT
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225.567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257.144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 343
; LENGTH: 5399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-343

Alignment Scores:
Pred. No.: 491      Length: 5399
Score: 48.00      Matches: 9
Percent Similarity: 90.91%      Conservative: 1
Best Local Similarity: 81.82%      Mismatches: 1
Query Match: 64.00%      Indels: 0
DB: 9      Gaps: 0
```

```

US-09-918-995-0770
: Sequence: 30770, Application US/09918995
: Publication No. US20030073623A1
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc.
: TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
: TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
: FILE REFERENCE: 20411-756
: CURRENT APPLICATION NUMBER: US/09/918, 995
: CURRENT FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: US/09/235, 076
: PRIOR FILING DATE: 1999-01-20
: NUMBER OF SEQ ID NOS: 38054
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 30770
: LENGTH: 263
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(263)
: OTHER INFORMATION: n = A,T,C or G
US-09-918-995-30770

Alignment Scores:
Pred. No.: 26 Length: 263
Score: 47.00 Matches: 9
Percent Similarity: 84.62% Conservative: 2
Best Local Similarity: 69.23% Mismatches: 2
Query Match: 62.67% Indels: 0
DB: 9 Gaps: 0

US-10-075-846-4_COPY_192-207 (1-16) x US-09-918-995-30770 (1-263)
Oy 3 SerLeucCySerProLeuProSerLeuSerLeuSer 15
Db 68 GCCTCCTCTGCTTCTTCTGCCCCCTCCTGCTCTTCT 106

RESULT 15
US-09-983-965-3940/C
: Sequence 3940, Application US/09983965
: Patent No. US20020137160A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengping
: APPLICANT: Byatt, John C.
: APPLICANT: Mathialagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 37-21(10297)C
: CURRENT APPLICATION NUMBER: US/09/983, 965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US 09/465, 231
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US 60/113, 678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 5912
: SEQ ID NO 3940
: LENGTH: 381
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: OTHER INFORMATION: Clone ID: 51-LIB3058-053-Q1-K1-E12
US-09-983-965-3940

Alignment Scores:
Pred. No.: 39 Length: 381
Score: 47.00 Matches: 9
Percent Similarity: 78.57% Conservative: 3
Best Local Similarity: 64.29% Mismatches: 2
Query Match: 62.67% Indels: 0
DB: 10 Gaps: 0

```



US-10-075-846-4\_COPY\_192\_207 (1-16) x US-09-983-965-3940 (1-381)

QY 2 SerSerIleLeuCySerProLeuProSerLeuSerLeuSer 15

Db 144 AGCGCATCTTGTGCTTACCGTCGTCCACAGCCTTGGAATCAGC 103

Search completed: July 1, 2003, 01:01:55  
Job time : 29.519 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus.p2n model

Run on: June 30, 2003, 23:00:29 ; Search time 84.1163 Seconds

(without alignments)  
3080.588 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSILCSPSLPSLSLV 16

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

## Command line parameters:

-MODL=frame\_plus.p2n.model -DEV=xlp  
-O/cgn2.1/USPTO.spool/US10075846/runat\_25062003.163648.5157/app.query.fasta\_1.782  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10075846.ecgn.1.1.3724.ernat.25062003.163648.5157 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOBJECT -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXt=7

## Database :

EST:\*  
1: em\_estba:\*  
2: em\_estchum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estlro:\*  
8: em\_hic:\*  
9: gb\_estl1:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	74.7	795	17 AG076357	AG076357 Pan trogl
2	55	73.3	310	14 B0484167	B0484167 pmp1c.pko
3	55	73.3	884	12 BE739001	BE739001 601556067
4	55	73.3	1081	12 BG622007	BG622007 602646601
5	54	72.0	128	17 A2485951	A2485951 1M0313112
6	53	70.7	263	10 BE526547	BE526547 M55F05STM
7	53	70.7	306	9 AA597350	AA597350 29621 lam
8	53	70.7	404	10 AV817296	AV817296 AV817296
9	53	70.7	470	10 BE591974	BE591974 WHE1651-1
10	53	70.7	547	10 AV548451	AV548451 AV548451
11	53	70.7	548	10 AV548438	AV548438 AV548438
12	53	70.7	556	10 AV547337	AV547337 AV547337
13	53	70.7	597	17 A0415379	A0415379 RPT-11-1
14	53	70.7	600	14 B0246138	B0246138 TAE15015F
15	53	70.7	615	13 B0268516	B0268516 B0268516
16	53	70.7	668	13 B0237354	B0237354 B0237354
17	53	70.7	787	17 A2198854	A2198854 SP_1038_B
18	53	70.7	1329	12 BG167249	BG167249 602342511
19	52	69.3	287	13 BM346474	BM346474 Kx71e05.y
20	52	69.3	364	9 A1164587	A1164587 A065P41U
21	52	69.3	443	17 A2713283	A2713283 RPT-24-6
22	52	69.3	445	13 B1449589	B1449589 dae73b10.
23	52	69.3	471	10 AW591971	AW591971 h136d10.x
24	52	69.3	484	17 A2826655	A2826655 256000.MA
25	52	69.3	510	12 BF430050	BF430050 256000.MA
26	52	69.3	537	17 A2966794	A2966794 2M0237P22
27	52	69.3	572	17 A2506723	A2506723 1M0348F03
28	52	69.3	656	17 AG041152	AG041152 Pan trogl
29	52	69.3	782	17 BH501609	BH501609 BOHLC72TR
30	52	69.3	834	13 B1091085	B1091085 602855289
31	51.5	68.7	654	12 BF691466	BF691466 602247635
32	51	68.0	231	9 AA967038	AA967038 ua39e01.r
33	51	68.0	242	12 BF552695	BF552695 UI-R-CO-1
34	51	68.0	303	10 AW346805	AW346805 29517 MAR
35	51	68.0	344	10 AW344669	AW344669 24424 MAR
36	51	68.0	359	10 BE117765	BE117765 UI-R-B51-
37	51	68.0	422	13 BM388008	BM388008 UI-R-CM1-
38	51	68.0	479	12 BG380127	BG380127 UI-R-CS0-
39	51	68.0	530	9 A1406681	A1406681 EST234968
40	51	68.0	540	13 B1896388	B1896388 fc80d10.x
41	51	68.0	549	10 BE117524	BE117524 UI-R-B51-
42	51	68.0	558	14 B0780785	B0780785 UI-R-FRO-
43	51	68.0	615	14 B0193942	B0193942 UI-R-CM1-
44	51	68.0	643	9 A1176553	A1176553 EST220140
45	51	68.0	650	17 AG053401	AG053401 Pan trogl

## ALIGNMENTS

RESULT 1  
AG076357/c 795 bp DNA linear GSS 03-NOV-2001  
LOCUS AG076357  
DEFINITION Pan troglodytes DNA, clone: PTB-070115.F, genomic survey sequence.  
ACCESSION AG076357  
VERSION AG076357.1 GI:16628159  
KEYWORDS GSS.  
SOURCE Pan troglodytes male lymphoblast DNA, clone: PTB Chimpanzee Male  
ORGANISM Pan troglodytes  
BAC library clone: PTB-070115.F.  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.  
REFERENCE  
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.

TITLE BAC end sequences of library PTB  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 795)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimpesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB. This BAC end  
was generated during the Rad process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1. 795  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-070115.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB chimpanzee Male BAC Library"  
BASE COUNT 201 a 115 c 325 g 149 t 5 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 765 Length: 795  
Score: 56.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 91.67% Mismatches: 0  
Query Match: 74.67% Indels: 0  
Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x AG076357 (1-795)  
QY 3 SerileuCySerProleupProserleuSerleu 14  
|||||  
Db 712 TCTATATTTATGTTCCCTCCCTCTCTCTCACGCTA 677  
RESULT 2 310 bp mRNA linear EST 07-JUN-2002  
LOCUS BQ484167  
DEFINITION pmpic.pk002.b9 Chicken Macrophage cDNA library Gallus gallus cDNA  
clone pmpic.pk002.b9 5', mRNA sequence.  
ACCESSION BQ484167  
VERSION BQ484167.1 GI:21328786  
KEYWORDS EST.  
SOURCE chicken.  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 310)  
Keeler, C.  
ESTs from Chicken Macrophages  
Unpublished (2002)  
Contact: Calvin Keeler  
University of Delaware  
Townsend Hall, Newark, DE 19717, USA  
Tel: 302-831-6473  
Fax: 302-831-2822  
Email: ckeeler@udel.edu  
FEATURES  
source  
1. 310  
/organism="Gallus gallus"  
/db\_xref="taxon:9031"  
/clone="pmpic.pk002.b9"

/clone\_lib="Chicken Macrophage cDNA library"  
/sex="Male and Female"  
/tissue\_type="LPS-stimulated HD-11 macrophage cells"  
/lab\_host="E.coli XL-10 Blue"  
/note="Vector: pBluescript II-SK"  
BASE COUNT 74 a 119 c 56 g 60 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 278 Length: 310  
Score: 55.00 Matches: 11  
Percent Similarity: 81.25% Conservative: 2  
Best Local Similarity: 68.75% Mismatches: 3  
Query Match: 73.33% Indels: 0  
Gaps: 0  
US-10-075-846-4\_COPY\_192\_207 (1-16) x BQ484167 (1-310)  
QY 1 SerSerleuCySerProleupProserleuSerleuSerVal 16  
|||||  
Db 196 TCTCTCTCAACCTGTGCTCCCTCCACACCTCCCTCCATC 243  
RESULT 3  
BE739001/c 884 bp mRNA linear EST 15-SEP-2000  
LOCUS BE739001  
DEFINITION 601556067F1 NIH\_MGC\_58 Homo sapiens cDNA clone IMAGE:3825809 5',  
mRNA sequence.  
ACCESSION BE739001  
VERSION BE739001.1 GI:10152993  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 884)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs.fda.gov  
Tissue Procurement: ATCC  
cDNA library Preparation: CLONETECH Laboratories, Inc.  
cDNA library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
http://image.llnl.gov  
Plate: L10M492 row: f column: 18  
High quality sequence stop: 555.  
FEATURES  
source  
1. 884  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3825809"  
/clone\_lib="NIH\_MGC\_58"  
/tissue\_type="hypertrophoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site\_1:  
SfiI (ggcgccctggcc); Site\_2: SfiI (ggcgccctggcc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGCGCGCGACATG-dT(30)BN-3'  
(where B = A, C, G, or G and N = A, C, G, or T). Average  
insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."  
BASE COUNT 223 a 208 c 280 g 173 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.22e+03 Length: 884

```

US-10-075-846-4_COPY_192_207 (1-16) x BE739001 (1-884)
Oy 4 1leucySerProleuProSerleuSerleuSer 15
Db 867 GTGCTGTGGCCTCTTCCTCCCTCCGCTTCCTTC 832

RESULT 4
LOCUS BG622007/c 1081 bp mRNA linear EST 18-APR-2001
DEFINITION BG622007 NIH_MGC_79 Homo sapiens CDNA clone IMAGE:4768033 5',
mRNA sequence.
ACCESSION BG622007
VERSION BG622007.1 GI:13673378
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM1631 row: b column: 02
High quality sequence stop: 205.
Location/Qualifiers
1. 1081
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4768033"
/clone_id="NIH_MGC_79"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site.1: Sfil (ggcgctcgccg); Site.2: Sfil (ggccatattggc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTCCTAGAGCGGCGCGCATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 527 a 194 c 325 g 35 t
ORIGIN

Alignment Scores:
Pred. No.: 1.62e+03 Length: 1081
Score: 55.00 Matches: 12
Percent Similarity: 86.67% Conservative: 1
Best Local Similarity: 80.00% Mismatches: 2
Query Match: 73.33% Indels: 0
DB: 12 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x BG622007 (1-1081)
Oy 1 SerSerleuLeucySerProleuProSerleuSerleuSer 15
Db 1076 TCTGCTTCGCTCTTCCTTCCTTCCTCCCTCCCTCTCTCTCT 1032

```

RESULT	5
AZ485951	
LOCUS	128 bp DNA linear GSS 05-OCT-2000
DEFINITION	M03J11R Mouse 10kb plasmid UOCCIM library Mus musculus genomic clone UOCCIM03J112 R, DNA sequence.
ACCESSION	AZ485951
VERSION	AZ485951.1 GI:10652245
KEYWORDS	GSS.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eumariola Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 128)
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamli,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished (2000)
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel.: 801 585 5606 Fax: 801 585 7177 Email: ddunne@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0313 Row: 1 Column: 12 Seq primer: CACACAGCAACACCTGTAACC Class: plasmid ends High quality sequence stop: 128. Location/Qualifiers 1..128 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UOCCIM03J112" /clone_lib="Mouse 10kb plasmid UOCCIM library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource ( <a href="http://www.jax.org/resources/documents/dnares/">http://www.jax.org/resources/documents/dnares/</a> ). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMDA2 (gii4732114[gb AF129072.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT	20 a 51 c 14 g 43 t
ORIGIN	
Alignment Scores:	
Pred. NO.:	110 Length: 128
Score:	54.00 Matches: 12
Percent Similarity:	86.67% Conservative: 1
Best Local Similarity:	80.00% Mismatches: 2
Query Match:	72.00% Indels: 0
DB:	17 Gaps: 0

```

US-10-075-846-4_COPY_192_207 (1-16) x AA597350 (1-128)
QY 2 SerSerIleucySerProleuProSerIleuSerIleuSerVal 16
||||| ||| |||||||||||||||||||||||:
Db 2 TCCTATTCTACTCTCTCCCTCCCTCTTCTTCTTCCCTG 46

RESULT 6
LOCUS BE526547/c 263 bp mRNA linear EST 19-MAR-2001
DEFINITION M55F0557M Arabidopsis developing seed Arabidopsis thaliana cDNA
ACCESSION BE526547
VERSION BE526547.1 GI:9784525
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 263)
REFERENCE White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
Ilarduya,O., Jaworski,J.G., Ohlrogge,J. and Banning,C.
A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
Plant Physiol. 124 (4), 1582-1594 (2000)
JOURNAL MEDLINE 20567808
COMMENT Contact: Banning, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI 48824
USA
Tel: 517 355 1609
Fax: 517 353 9334
Email: benningsmsu.edu
Clones were originally prepared at Michigan State University.
Arabidopsis Biological Resource Center, The Ohio State University,
309 Botany & Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210
USA, FAX: 6142920603 TEL: 6142929371.
Location/Qualifiers
FEATURES
source
1. .263
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="600035219R1"
/clone_lib="Arabidopsis developing seed"
/tissue_type="seed"
/dev_stage="5-13 days after flowering"
/lab_host="E.coli"
/note="Organ: Developing seed; Vector: pBluescript SK-;
Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 85 a 38 c 79 g 61 t
ORIGIN

Alignment Scores:
Pred. No.: 416 Length: 263
Score: 53.00 Matches: 10
Percent Similarity: 81.258 Conservative: 3
Best Local Similarity: 62.508 Mismatches: 3
Query Match: 70.678 Indels: 0
DB: AV817296 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x BE526547 (1-263)
QY 1 SerSerIleucySerProleuProSerIleuSerIleuSerVal 16
||||| ||| |||||||||||||||||||||||:
Db 130 TCATTATCAGTAGCTTCTCCACCTCTCCCTCAATCTCGTATCACTT 83

RESULT 7
LOCUS AA597350 306 bp mRNA linear EST 19-SEP-1997
DEFINITION 29621 Lambda-PRL2 Arabidopsis thaliana cDNA clone 196B17M4, mRNA
sequence.
ACCESSION AA597350
AUTHORS
```

```

VERSION AA597350.1 GI:2412773
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 306)
REFERENCE Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
JOURNAL MEDLINE 95148729
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22333ctn@lm.cl.msu.edu
Seq primer: 5' end dye terminator primer closer than T7.
Location/Qualifiers
FEATURES
source
1. .306
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="196B17M4"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Z1p-lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
siliques). The vector is BRL's lambda Z1p-lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA.
BASE COUNT 97 a 40 c 82 g 82 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 515 Length: 306
Score: 53.00 Matches: 10
Percent Similarity: 81.258 Conservative: 3
Best Local Similarity: 62.508 Mismatches: 3
Query Match: 70.678 Indels: 0
DB: AV817296 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AA597350 (1-306)
QY 1 SerSerIleucySerProleuProSerIleuSerIleuSerVal 16
||||| ||| |||||||||||||||||||||||:
Db 79 TCATTATCAGTAGCTTCTCCACCTCTCCCTCAATCTCGTATCACTT 32

RESULT 8
LOCUS AV817296 404 bp mRNA linear EST 01-APR-2002
DEFINITION AV817296 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-93-O12 3',
mRNA sequence.
ACCESSION AV817296
VERSION AV817296.1 GI:19859167
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 404)
REFERENCE Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
```



```

US-10-075-846-4_COPY_192_207 (1-16) x AV548451 (1-547)
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTATCAGTACGCTGTCCACCTCTCCCTTCATCTCTGTATCAGTT 406

RESULT 11
AV548438 548 bp mRNA linear EST 06-SEP-2000
AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RFL54C06F 3', mRNA sequence.
AV548438
AV548438.1 GI:8719851
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Source
1..548
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZL54C06F"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/issue="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 176 a 131 c 84 g 157 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548438 (1-548)
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTATCAGTACGCTGTCCACCTCTCCCTTCATCTCTGTATCAGTT 406

RESULT 11
AV548438 548 bp mRNA linear EST 06-SEP-2000
AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RFL54C06F 3', mRNA sequence.
AV548438
AV548438.1 GI:8719851
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Source
1..548
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZL54C06F"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/issue="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 176 a 131 c 84 g 157 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 548
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548438 (1-548)
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTATCAGTACGCTGTCCACCTCTCCCTTCATCTCTGTATCAGTT 406

RESULT 11
AV548438 548 bp mRNA linear EST 06-SEP-2000
AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RFL54C06F 3', mRNA sequence.
AV548438
AV548438.1 GI:8719851
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Source
1..548
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZL54C06F"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/issue="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 175 a 131 c 81 g 160 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548438 (1-548)
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTATCAGTACGCTGTCCACCTCTCCCTTCATCTCTGTATCAGTT 406

RESULT 11
AV548438 548 bp mRNA linear EST 06-SEP-2000
AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RFL54C06F 3', mRNA sequence.
AV548438
AV548438.1 GI:8719851
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Source
1..548
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZL54C06F"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/issue="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 175 a 131 c 81 g 160 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 547
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548438 (1-548)
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTATCAGTACGCTGTCCACCTCTCCCTTCATCTCTGTATCAGTT 406

RESULT 11
AV548438 548 bp mRNA linear EST 06-SEP-2000
AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RFL54C06F 3', mRNA sequence.
AV548438
AV548438.1 GI:8719851
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093

JOURNAL MEDLINE
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
Source
1..548
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="RZL54C06F"
/clone_1lb="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/issue="Vector: pluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 176 a 131 c 84 g 157 t
ORIGIN

Alignment Scores:
Pred. No.: 1.17e+03 Length: 548
Score: 53.00 Matches: 10
Percent Similarity: 81.25% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 3
Query Match: 70.67% Indels: 0
DB: 10 Gaps: 0

US-10-075-846-4_COPY_192_207 (1-16) x AV548438 (1-548)
Oy 1 SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal 16
    ||| |||:: ||| ||||| |||||::|||::||| |||
    359 TCTATTATCAGTACGCTGTCCACCTCTCCCTTCATCTCTGTATCAGTT 406

RESULT 11
AV548438 548 bp mRNA linear EST 06-SEP-2000
AV548438 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RFL54C06F 3', mRNA sequence.
AV548438
AV548438.1 GI:8719851
EST.
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7
```

```

DB          396 TCTATTACGTAAGCTTGTCCACACTCTCCCAATCATCTCTGATATCAGTT 443
RESULT 12
LOCUS       AV547337                      556 bp    mRNA    Linear   EST 06-SEP-2000
DEFINITION  AV547337 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
ACCESSION   AV547337
VERSION     AV547337.1 GI:8718751
KEYWORDS
SOURCE      EST.
            thale cress.
ORGANISM    Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE   1 (bases 1 to 556)
AUTHORS     Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
TITLE       A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
JOURNAL     DNA Res. 7, 175-180 (2000)
MEDLINE     20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
            source
                1..556
                    /organism="Arabidopsis thaliana"
                    /strain="Columbia"
                    /db_xref="taxon:3702"
                    /clone="RZL29h10F"
                    /clone_lib="Arabidopsis thaliana roots Columbia"
                    /tissue_type="roots"
                    /note="Vector: pBluescriptII SK-, Site_1: EcoRI; Site_2:
                        XhoI"
BASE COUNT  145 a      144 c      88 g      179 t
ORIGIN
Alignment Scores:
Score: No.:           1.2e+03      Length:           556
Percd:              53.00         Matches:             10
Best Local Similarity:  81.25%     Conservative:        3
Percent Similarity:    62.50%     Mismatches:          0
Query Match:          70.67%     Indels:               0
DB:                   10         Gaps:                 0
US-10-075-846-4_COPY_192_207 (1-16) x AV547337 (1-556)
Oy          1 SerSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerVal 16
            ||| |:::||||| ||||| ||||| ||||| ||||| ||||| |||||
Ddb         221 TCTATTACGTAAGCTTGTCCACACTCTCCCAATCATCTCTGATATCAGTT 268
RESULT 13
LOCUS       AQ415379                      597 bp    DNA    linear   GSS 23-MAR-1999
DEFINITION  AQ415379 RPCI-11 Homo sapiens genomic clone RPCI-11-177K5,
ACCESSION   AQ415379
VERSION     AQ415379.1 GI:4474348
KEYWORDS    GSS.
SOURCE      human.
            Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE   1 (bases 1 to 597)
AUTHORS     Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
            J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
            Unpublished (1997)

```



**COMMENT**

Other GSSS: RPCI-11-177K5.TV  
Contact: Shaying Zhao, William Niernan, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeet@igr.org  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics Inc. ([info@resgen.com](http://info@resgen.com)). BAC end search page: [http://www.tlgr.org/tldb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tlgr.org/tldb/humgen/bac_end_search/bac_end_search.html).  
Seq primer: SP6  
Class: BAC ends.

**FEATURES**

**source**

location/Qualifiers

1..597

/organism="Homo sapiens"  
/db\_xref="GDB:7567828".  
/db\_xref="taxon:9606"  
/clone="RPCI-11-177K5"  
/clone\_lib="RPCI-11"  
/sex="Male"  
/cell\_type="lymphocytes"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPCI11 Human Male BAC Library"

**BAC COUNT**

142 a 163 c 127 g 164 t 1 others

**BASE COUNT**

Alignment Scores:

Pred. No.: 1.32e+03 Length: 597

Score: 53..00 Matches: 9

Percent Similarity: 92.86% Conservative: 4

Best local Similarity: 64.29% Mismatches: 1

Query Match: 70.67% Indels: 0

DB: 17 Gaps: 0

US-10-075-846-4\_COPY\_192-207 (1-16) x A0415379 (1-597)

Oy 2 SerSerIleucysSerProLeuProSerIleuSerIeuSer 15  
|||||:||||| |||||:::|||||

Db 105 AGTCTGTGGATGTGGCCCTCCTCCCAACATGGCACATTTC 146

**RESULT 14**

B0246138/c 600 bp mRNA linear EST\_03-MAY-2002

**LOCUS**

Definition TaEI5 Triticum aestivum cDNA clone TaEI5015F1OR, mRNA sequence.  
Accession B0246138  
Version B0246138  
Keywords GI:20442014  
Source EST.  
Organism bread wheat.  
Triticum aestivum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
1 (bases 1 to 600)  
Cloutier.S.  
Wheat functional genomics - Glenlea developing seeds cDNA libraries  
Unpublished (2002)  
Contact: Dr. Sylvie Cloutier  
Cereal Research Centre, Agriculture and Agri-Food Canada  
195 Dfoce Rd, Winnipeg, MB, Canada R3T 2M9  
Tel: (204) 983-2340  
Fax: (204) 983-4604  
Email: scloutier@em.agr.ca  
was cloned directionally, not all sequences generated with reverse primer were from the 5' end (same with forward primer and 3' end).  
Average insert size is >1.4 kb  
plate: 015 row: F column: 10  
Seq primer: M13 Reverse.  
Location/Qualifiers

**FEATURES**

```

source
1. .600
/organism="Triticum aestivum"
/cultivar="Glenlea"
/db_xref="taxon:4565"
/clone_1b="TAE15015F10R"
/clone_1b="TAE15"
/tissue_type="developing seeds"
/dev_stage="15 days after anthesis"
/lab_host="E. coli DH10B"
/note="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
Site_1: NotI; Site_2: MuiI; mRNA obtained from wheat seeds
of cultivar Glenlea 15 days post-anthesis"

BASE COUNT      185 a      126 c      183 g      106 t

Alignment Scores:
Pred. NO.:      1.33e+03      Length:      600
Score:      53.00      Matches:      11
Percent Similarity:      75.00%      Mismatches:      4
Best Local Similarity:      68.75%      Indels:      0
Query Match:      70.67%      Gaps:      0
14

US-10-075-846-4_COPY_192_207 (1-16) x BQ246138 (1-600)

Oy      1 SerSerSer1LeuCySSerProLeuProSerLeuSerLeuSerVal 16
Db      67 AGCAGCTCTTCTGCTGACCCCTGCTCCCTGTTGAGCAGACGCTT 20
|||||||      |||||||||      |||||      |||::|

RESULT 15
LOCUS      BJ268516      615 bp      mRNA      linear      EST 09-APR-2002
DEFINITION      BJ268516 Y. Ogihara unpublished cDNA library, wh_oh Triticum
aestivum cDNA clone wh0h17f17 5', mRNA sequence.
ACCESSION      BJ268516
VERSION      BJ268516.1
KEYWORDS      GI:20088515
SOURCE      EST.
ORGANISM      bread wheat.
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Triticum:
1 (bases 1 to 615)
Ogihara,Y. and Mural,K.
Expressed genes in Triticum aestivum
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information.
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .615
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="wh0h17f17"
/clone_1b="Y. Ogihara unpublished cDNA library, wh_oh"
/tissue_type="pistill at heading date"
/dev_stage="Feekes" scale 10.5"
/note="Vector: Lambda Uni-ZAP XR, excised phagmid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown under
hydroponic conditions at UC Davis, salt stressed for 12
hours, and for 7 days, then dissected and frozen (Akhunov
in J Dvorak Lab). Total RNA was prepared from sheath
tissue, equal quantities of RNA were pooled from the two
samples, polyA was purified from the pooled RNA, a cDNA
library was made, and the cDNA clones were in vivo
excised to give pluscript phagemids in the T7 Close lab
at the University of California, Riverside (Akhunov, Chirp
, Choi, Close, Fenton, Kianian, Otto, Simons, Zhang).

```

Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors). "

BASE COUNT	184 a	127 c	195 g	108 t	1 others
ORIGIN					

Alignment Scores:

Pred. No.:	1.38e+03	Length:	615
Score:	53.00	Matches:	11
Percent Similarity:	75.00%	Conservative:	1
Best Local Similarity:	68.75%	Mismatches:	4
Query Match:	70.67%	Indels:	0
DB:	13	Gaps:	0

US-10-075-846-4\_COPY\_192\_207 (1-16) x BJ268516 (1-615)

QY	1	SerSerSerIleLeuCySerProLeuProSerLeuSerLeuSerVal	16
Db	189	AGCAGCTCTTTCGCTGCTAGCCCTCTCCCTGTTGAGCAGCAGCCTT	142

Search completed: July 1, 2003, 00:55:32  
Job time : 89.1163 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:14 ; Search time 2.32662 Seconds  
(without alignments)  
916.354 Million cell updates/sec

```
Title: US-10-075-846-4_COPY_192_207
Perfect score: 75
Sequence: 1 SSSIICSPPLSLSLSV 16
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```
Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

A\_Geneseq\_101002.\*

1	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT *
2	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT *
3	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT *
4	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT *
5	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT *
6	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT *
7	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT *
8	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT *
9	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT *
10	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT *
11	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT *
12	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT *
13	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT *
14	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT *
15	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT *
16	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT *
17	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT *
18	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT *
19	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT *
20	/SDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT *
21	/SDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT *
22	/SDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT *
23	/SDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	49	65.3	826	21	AA038474	Arabidopsis thaliana
2	49	65.3	855	21	AA038473	Arabidopsis thaliana
3	49	65.3	888	21	AA038472	Arabidopsis thaliana
4	47	62.7	83	22	AA094754	Human reproductive
5	47	62.7	160	22	AA094752	Human protein sequ
6	46	61.3	490	23	AB092207	Hericidially active
7	45.5	60.7	67	22	AB023900	Novel human diagno
8	45.5	60.7	89	22	AB011062	Novel human diagno
9	45	60.0	166	22	AA023535	HHV-2 EST encoded
10	45	60.0	166	22	AA014387	Human novel protei

11	58.7	75	20	AAV14488	Fragment of human
12	58.7	120	20	ABB17174	Human nervous syst
13	58.7	149	21	AAAG00176	Human secreted pro
14	58.7	887	22	AAAM39431	Human polypeptide
15	57.3	274	22	AAAM39109	Human polypeptide
16	57.3	289	22	AAAG40895	Human polypeptide
17	57.3	3554	23	AAAE20147	Mouse C3b/C4b comp
18	56.0	39	22	ABBI5852	Human nervous syst
19	56.0	739	19	AAAM61532	Mouse Fas-binding
20	56.0	739	22	AAAB69150	Mouse daxx protein
21	55.3	73	22	AAAM89900	Human immune/haema
22	54.7	50	22	AAU60466	Proptiomibacterium
23	54.7	69	22	AAU62995	Proptiomibacterium
24	54.7	70	23	ABP05968	Human ORFX protein
25	54.7	73	22	AAU59452	Proptiomibacterium
26	54.7	109	22	AAAM89167	Human immune/haema
27	54.7	116	22	AAU44557	Proptiomibacterium
28	54.7	215	22	AAAG98446	Gorilla olfactory
29	54.7	223	20	AAV74114	Human prostate tum
30	54.7	538	23	AAO21708	Human secreted .pro
31	53.3	60	22	AAU33291	Novel human secret
32	53.3	68	23	ABP05420	Human ORFX protein
33	53.3	72	22	AAU46650	Proptiomibacterium
34	53.3	93	22	AAU49188-	Proptiomibacterium
35	53.3	98	21	AAV73448	Human secreted pro
36	53.3	106	21	AAAB40561	Human ORFX ORF325
37	53.3	133	23	AAO17460	Human liver cancer
38	53.3	141	21	AAV54348-	Amino acid sequenc
39	53.3	209	22	ABG04102	Novel human diagno
40	53.3	262	20	AAV37102	Chlamydia trachoma
41	53.3	265	22	AAU16053	Human novel secret
42	53.3	296	22	AAAG24854	Arabidopsis thailia
43	53.3	301	21	AAAG24853	Arabidopsis thailia
44	53.3	353	22	AAAG16162	Novel human diagno
45	53.3	395	21	AAAG42022	Arabidopsis thailia

## ALIGNMENTS

XX	RESULT 1
AAAG38474	
ID	AAAG38474 standard; Protein: 826 AA.
XX	
AC	AAAG38474;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 47470.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123160.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141642.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-01452913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 09-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160960.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 65.3%; Score 49; DB 21; Length 826;  
Best Local Similarity 73.3%; Pred. No. 19;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSILCPLPSLSLS 15  
Db 431 SSSILMDVPPLSLSS 445

## RESULT 2

AAG38473  
ID AAG38473 standard; Protein; 855 AA.

XX AAG38473;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 47469.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123160.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 23-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0127462.

XX 01-APR-1999; 99US-0128234.

XX 06-APR-1999; 99US-0128714.

XX 08-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.



PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.

PR 25-OCT-1999: 99US-0161405.  
PR 25-OCT-1999: 99US-0161406.  
PR 26-OCT-1999: 99US-0161359.  
PR 26-OCT-1999: 99US-0161360.  
PR 26-OCT-1999: 99US-0161361.  
PR 28-OCT-1999: 99US-0161920.  
PR 28-OCT-1999: 99US-0161992.  
PR 28-OCT-1999: 99US-0161993.  
PR 29-OCT-1999: 99US-0162142.

Query Match 65.3%; Score 49; DB 21: Length 888;  
Best Local Similarity 73.3%; Pred. No. 20;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLIS 15  
11111 1:11111  
Db 493 SSSIIMDPYPLSLIS 507

RESULT 4  
AAM94754  
ID AAM94754 standard: Protein; 83 AA.  
XX  
AC AAM94754;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human reproductive system related antigen SEQ ID NO: 3412.  
XX  
KW Human; reproductive system related antigen; reproductive system disorder;  
cancer; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01339.  
XX  
PR 31-JAN-2000: 2000US-0179065.  
PR 04-FEB-2000: 2000US-0180628.  
PR 24-FEB-2000: 2000US-0184664.  
PR 02-MAR-2000: 2000US-0186350.  
PR 16-MAR-2000: 2000US-019874.  
PR 17-MAR-2000: 2000US-0190076.  
PR 18-APR-2000: 2000US-0198123.  
PR 19-MAY-2000: 2000US-0205515.  
PR 07-JUN-2000: 2000US-0209467.  
PR 28-JUN-2000: 2000US-0214886.  
PR 30-JUN-2000: 2000US-0215135.  
PR 07-JUL-2000: 2000US-0216647.  
PR 07-JUL-2000: 2000US-0216880.  
PR 11-JUL-2000: 2000US-0217487.  
PR 11-JUL-2000: 2000US-0217496.  
PR 14-JUL-2000: 2000US-0218290.  
PR 26-JUL-2000: 2000US-0220963.  
PR 26-JUL-2000: 2000US-0220964.  
PR 14-AUG-2000: 2000US-0224518.  
PR 14-AUG-2000: 2000US-0224519.  
PR 14-AUG-2000: 2000US-0225213.  
PR 14-AUG-2000: 2000US-0225214.  
PR 14-AUG-2000: 2000US-0225266.  
PR 14-AUG-2000: 2000US-0225267.  
PR 14-AUG-2000: 2000US-0225268.  
PR 14-AUG-2000: 2000US-0225270.  
PR 14-AUG-2000: 2000US-0225447.  
PR 14-AUG-2000: 2000US-0225757.  
PR 14-AUG-2000: 2000US-0225758.  
PR 14-AUG-2000: 2000US-0225759.  
PR 18-AUG-2000: 2000US-0226279.  
PR 22-AUG-2000: 2000US-0226681.  
PR 22-AUG-2000: 2000US-0226686.

PR 22-AUG-2000: 2000US-0227182.  
PR 23-AUG-2000: 2000US-0227009.  
PR 30-AUG-2000: 2000US-0228924.  
PR 01-SEP-2000: 2000US-0229287.  
PR 01-SEP-2000: 2000US-0229343.  
PR 01-SEP-2000: 2000US-0229344.  
PR 01-SEP-2000: 2000US-0229345.  
PR 05-SEP-2000: 2000US-0229509.  
PR 05-SEP-2000: 2000US-0229513.  
PR 06-SEP-2000: 2000US-0230437.  
PR 06-SEP-2000: 2000US-0230438.  
PR 08-SEP-2000: 2000US-0231242.  
PR 08-SEP-2000: 2000US-0231243.  
PR 08-SEP-2000: 2000US-0231244.  
PR 08-SEP-2000: 2000US-0231413.  
PR 08-SEP-2000: 2000US-0231414.  
PR 08-SEP-2000: 2000US-0232080.  
PR 08-SEP-2000: 2000US-0232081.  
PR 12-SEP-2000: 2000US-0231968.  
PR 14-SEP-2000: 2000US-0232397.  
PR 14-SEP-2000: 2000US-0232398.  
PR 14-SEP-2000: 2000US-0232399.  
PR 14-SEP-2000: 2000US-0232400.  
PR 14-SEP-2000: 2000US-0232401.  
PR 14-SEP-2000: 2000US-0233063.  
PR 14-SEP-2000: 2000US-0233064.  
PR 14-SEP-2000: 2000US-0233065.  
PR 21-SEP-2000: 2000US-0234223.  
PR 21-SEP-2000: 2000US-0234274.  
PR 25-SEP-2000: 2000US-0234997.  
PR 25-SEP-2000: 2000US-0234998.  
PR 26-SEP-2000: 2000US-0235484.  
PR 27-SEP-2000: 2000US-0235634.  
PR 27-SEP-2000: 2000US-0235636.  
PR 29-SEP-2000: 2000US-0236327.  
PR 29-SEP-2000: 2000US-0236367.  
PR 29-SEP-2000: 2000US-0236368.  
PR 29-SEP-2000: 2000US-0236369.  
PR 29-SEP-2000: 2000US-0236370.  
PR 02-OCT-2000: 2000US-0236802.  
PR 02-OCT-2000: 2000US-0237037.  
PR 02-OCT-2000: 2000US-0237038.  
PR 02-OCT-2000: 2000US-0237039.  
PR 02-OCT-2000: 2000US-0237040.  
PR 13-OCT-2000: 2000US-0239935.  
PR 13-OCT-2000: 2000US-0239937.  
PR 20-OCT-2000: 2000US-0240960.  
PR 20-OCT-2000: 2000US-0241221.  
PR 20-OCT-2000: 2000US-0241785.  
PR 20-OCT-2000: 2000US-0241786.  
PR 20-OCT-2000: 2000US-0241787.  
PR 20-OCT-2000: 2000US-0241808.  
PR 20-OCT-2000: 2000US-0241809.  
PR 20-OCT-2000: 2000US-0241826.  
PR 01-NOV-2000: 2000US-0244617.  
PR 01-NOV-2000: 2000US-0244674.  
PR 08-NOV-2000: 2000US-0246475.  
PR 08-NOV-2000: 2000US-0246476.  
PR 08-NOV-2000: 2000US-0246477.  
PR 08-NOV-2000: 2000US-0246478.  
PR 08-NOV-2000: 2000US-0246523.  
PR 08-NOV-2000: 2000US-0246524.  
PR 08-NOV-2000: 2000US-0246525.  
PR 08-NOV-2000: 2000US-0246526.  
PR 08-NOV-2000: 2000US-0246527.  
PR 08-NOV-2000: 2000US-0246528.  
PR 08-NOV-2000: 2000US-0246532.  
PR 08-NOV-2000: 2000US-0246609.  
PR 08-NOV-2000: 2000US-0246610.  
PR 08-NOV-2000: 2000US-0246611.  
PR 08-NOV-2000: 2000US-0246613.  
PR 17-NOV-2000: 2000US-0249207.  
PR 17-NOV-2000: 2000US-0249208.





XX WO200210210-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 28-AUG-2001; 2001WO-EP09892.  
XX  
XX 28-AUG-2001; 2001WO-EP09892.  
XX  
XX (FARB ) BAYER AG.  
XX  
XX Tietjen K, Weidler M;  
XX  
XX WPI: 2002-269010/31.  
XX  
XX  
XX Identifying plant target proteins for herbicidically active compounds,  
XX comprising aligning and comparing nucleic acid or amino acid sequences  
XX from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms -  
XX  
XX Claim 5; SEQ ID NO 1418; 261pp + Sequence Listing; English.  
XX  
XX The invention relates to identifying target proteins  
XX (AB80790-AB894016) for herbicidically active compounds, comprising  
XX aligning and comparing nucleic acid or amino acid sequences from plant  
XX with nucleic acid or amino acid sequences from non-plant organisms using  
XX suitable search parameters, where plant sequences having an E-value  
XX greater by a factor of 3 than the E-value of most similar non-plant  
XX sequences are selected. The polypeptides or nucleic acids encoding them  
XX are useful for identifying modulators. The identified modulators are  
XX useful as herbicides.  
XX  
XX  
XX Sequence 490 AA:  
XX  
XX Query Match 61.3%; Score 46; DB 23; Length 490;  
XX Best Local Similarity 68.8%; Pred. No. 33;  
XX Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
XX  
XX QY 1 SSSIICSPLPSPSLV 16  
XX :|||:|||||:|  
XX Db 210 SSSILSLPLPPSLPV 225  
XX  
XX  
XX RESULT 7  
XX ABG23990  
XX ID ABG23990 standard; Protein; 67 AA.  
XX  
XX AC ABG23990;  
XX  
XX DT 18-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #23981.  
XX  
XX KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX  
XX PR 31-MAR-2000; 2000US-0540217.  
XX  
XX PR 23-AUG-2000; 2000US-0649167.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
XX  
XX N-PSDB; AAS8177.  
XX  
XX

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity -  
XX  
XX Claim 20; SEQ ID No 54349; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
XX Sequence 67 AA:  
XX  
XX Query Match 60.7%; Score 45.5; DB 22; Length 67;  
XX Best Local Similarity 78.6%; Pred. No. 4.9;  
XX Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
XX  
XX QY 1 SSSIICSPLPSPSL 14  
XX :|||:|||||:|  
XX Db 20 TSFSLCPLP-LSL 32  
XX  
XX  
XX RESULT 8  
XX ABG11062  
XX ID ABG11062 standard; Protein; 89 AA.  
XX  
XX AC ABG11062;  
XX  
XX DT 18-FEB-2002 (first entry)  
XX  
XX DE Novel human diagnostic protein #11053.  
XX  
XX KW Human: chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200175067-A2.  
XX  
XX PD 11-OCT-2001.  
XX  
XX PF 30-MAR-2001; 2001WO-US08631.  
XX  
XX PR 31-MAR-2000; 2000US-0540217.  
XX  
XX PR 23-AUG-2000; 2000US-0649167.  
XX  
XX PA (HYSE-) HYSEQ INC.  
XX  
XX PI Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI: 2001-639362/73.  
XX  
XX N-PSDB; AAS75249.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX

PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
PS Claim 20; SEQ ID No 41421; 103bp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 89 AA:  
Query Match 60.7%; Score 45.5; DB 22; Length 89;  
Best Local Similarity 78.6%; Pred. No. 6.6;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 SSITCSPLPSLSL 14  
: 111111111111  
Db 49 TSSFLCSPLP-LSL 61  
RESULT 9  
AAM23635  
ID AAM23635 standard; Protein: 166 AA.  
AC AAM23635;  
XX  
XX 12-OCT-2001 (first entry)  
DT  
XX  
DE HHV-2 EST encoded protein SEQ ID NO: 1160.  
XX  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.  
XX  
OS Human herpesvirus 2.  
XX  
XX WO200154477-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 25-JAN-2001; 2001WO-US02687.  
XX  
XX 25-JAN-2001; 2000US-0491404.  
XX 25-JAN-2001; 2000US-0617746.  
XX 17-JUL-2000; 2000US-0631451.  
XX 03-AUG-2000; 2000US-0631451.  
XX 15-SEP-2000; 2000US-0663870.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX  
XX WPI: 2001-476164/51.  
DR

DR N-PSDB; AAM98294.  
XX  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX  
XX Claim 20; Page 850-851; 1275bp; English.  
PS  
XX  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.  
XX  
SQ Sequence 166 AA:  
Query Match 60.0%; Score 45; DB 22; Length 166;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 5 LCSPPLPSLSLV 16  
: 111111111111  
Db 143 LCPPLPSLSCAI 154  
RESULT 10  
AAU14387  
ID AAU14387 standard; Protein: 166 AA.  
AC AAU14387;  
XX  
XX 24-OCT-2001 (first entry)  
DT  
XX  
DE Human novel protein #258.  
XX  
XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;  
KW immunomodulatory; cyostatic; neuroprotective; vulnerrary; nootropic;  
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;  
KW antibacterial; antiallergic; dermatological; hemostatic; antiasthmatic;  
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;  
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;  
KW tissue regeneration; immune disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO200154437-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 25-JAN-2001; 2001WO-US02623.  
XX  
XX 25-JAN-2000; 2000US-0491404.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI: 2001-451939/48.  
XX DR N-PSDB; AAS22692.  
XX  
XX Isolated polypeptides useful for treating anti-inflammatory diseases,  
PT nervous system disorders, and for regenerating bone and cartilage -  
XX  
XX Example 4; Page 795-796; 894bp; English.  
XX  
XX The invention relates to polynucleotides encoding novel human  
CC proteins or their active domains. The polypeptides, polynucleotides and  
CC antibodies raised against the polypeptides are used in a method of  
CC treatment of a mammal and prevention of disorders caused by the aberrant  
CC protein expression or activity. The polypeptides can be used as  
CC molecular weight markers, food supplements, and in antibody production.  
CC

CC The polypeptides are used to identify compounds which bind to the  
CC polypeptides. Polynucleotides of the invention are used as probes and  
CC primers, for sequencing, for chromosome or gene mapping, in the  
CC production of recombinant proteins, and in generating anti-sense DNA or  
CC RNA and in gene therapy. Polypeptides of the invention can be used to  
CC target drugs to a tumour. In assays to determine biological activity, to  
CC raise antibodies/ elicit an immune response, to determine quantitative  
CC protein levels, as tissue markers, and to isolate receptors or ligands.  
CC Polypeptides of the invention may also be useful in treating platelet  
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,  
CC ligament and/or nerve tissue, wound healing, treating burns, promoting  
CC the proliferation, differentiation and survival of stem cells, as a  
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,  
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral  
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or  
CC fungal infection or from autoimmunity cancer, allergy, asthma,  
CC graft-versus-host disease, eczema, haemophilia, thrombosis,  
CC anti-inflammatory diseases, nervous system disorders, and infection.  
CC The present sequence represents a protein of the invention.

SQ Sequence 166 AA;

Query Match 60.0%; Score 45; DB 22; Length 166;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 LCSPLPSLSLV 16  
11 ||||| :  
Db 143 LCSPLPSLSLCAI 154

RESULT 11  
AA14488

ID AA14488 standard; Protein: 75 AA.

XX AA14488;

DT 17-AUG-1999 (first entry)

XX Fragment of human secreted protein encoded by gene 33.

XX Human: secreted protein; fusion protein; gene therapy; protein therapy;  
XX diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
XX developmental abnormality; foetal deficiency; blood; allergy; renal;  
XX immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
XX inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
XX cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
XX osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
XX endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX Homo sapiens.

PN MO9919339-A1.

XX 22-APR-1999.

XX 08-OCT-1998; 98WO-US21142.

XX 09-OCT-1997; 97US-0071498.

XX 09-OCT-1997; 97US-0061463.

XX 09-OCT-1997; 97US-0061527.

XX 09-OCT-1997; 97US-0061529.

XX 09-OCT-1997; 97US-0061532.

XX 09-OCT-1997; 97US-0061536.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Duan R, Ebner R, Ferrie AM, Florence C;

PI Florence KA, Greene JM, Olsen HS, Rosen CA, Ruben SM;

XX Young PE, Yu G;

XX WPI; 1999-277587/23.

PT New isolated human genes and the secreted polypeptides they encode  
XX  
XX  
PS Disclosure; Page 45; 226pp; English.

CC This sequence represents a fragment of a secreted human protein encoded  
CC by the nucleic acid molecule detailed in the descriptor line. The gene  
CC can be used to generate fusion proteins by linking to the gene to a  
CC human immunoglobulin Fc portion (e.g. AAX79002) for increasing the  
CC stability of the fused protein as compared to the human protein only.  
CC The invention relates to 53 novel genes and their fragments (nucleic  
CC acid sequences: AAX79011-X79064; amino acid sequences AA14411-Y14464)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 53  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAX79011 for described uses).

SQ Sequence 75 AA;

Query Match 58.7%; Score 44; DB 20; Length 75;  
Best Local Similarity 66.7%; Pred. No. 9.5;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSILCSPLPSLS 13  
11 ||||| :  
Db 15 TSQLCSPMSPLN 26

RESULT 12  
ABB17174

ID ABB17174 standard; Protein: 120 AA.

XX ABB17174;

DT 23-JAN-2002 (first entry)

XX Human nervous system related polypeptide SEQ ID NO 5831.

XX Human: nootropic; neuroprotective; cytosolic; dermatological; virocidic;  
XX immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;  
XX antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;  
XX antihemagic; hepatotropic; cerebroprotective; antiinflammatory;  
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
XX antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
XX neurological disease; infection; nephrotropic; gene therapy; vaccine.

XX Homo sapiens.

PN MO200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US01334.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205415.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 03-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231966.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249224.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251866.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPL: 2001-541565/60.  
N-PSDB: ABA13500.  
  
Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
useful for preventing, diagnosing and/or treating nervous system  
cancers and metastases -  
  
Claim 11: SEQ ID NO 5831: 1701pp + Sequence Listing: English.

The invention relates to novel genes (ABA11004-ABA21534) and proteins (ABA14678-ABA18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; and (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and



Best Local Similarity 69.2%; Pred. No. 1.3e+02;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 SSSILCSPPLPSLS 13  
|||:|||||:|  
Db 442 SSOITTSPLPSVS 454

## RESULT 15

AAM39109

ID AAM39109 standard; Protein: 274 AA.

AC AAM39109;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 2254.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0596042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 03-AUG-2000; 2000US-0653450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.

PA (HYSE-) HYSEQ INC.

PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao Q, Zhou P, Goodrich R, Drmanec RT;

XX WPI: 2001-442253/47.  
DR N-PSDB: AA158265.

PT Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries.

PS Example 4; SEQ ID NO 2254; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: immune system suppression,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.

SQ Sequence 274 AA;

Query Match

57.3%; Score 43; DB 22; Length 274;

Best Local Similarity

71.4%; Pred. No. 54;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 SSSILCSPPLPSLSL 14

Db 145 SPSLQCPPLPSLSL 158

Search completed: June 25, 2003, 17:15:33  
Job time : 3.32662 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:14:25 ; Search time 0.85906 Seconds

(without alignments)  
548.002 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Sequence: 1 SSILCSPLPSLSLV 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext. 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/Dackfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	44	58.7	75	4	US-09-288-143-205
2	43	57.3	106	4	US-09-177-249-208
3	42	56.0	739	4	US-09-022-983-2
4	39	52.0	291	4	US-09-328-869-2
5	39	52.0	291	4	US-09-629-774A-2
6	38	50.7	13	1	US-08-299-187-13
7	38	50.7	13	5	PCT-US95-11114-13
8	38	50.7	251	1	US-07-956-700B-84
9	38	50.7	251	1	US-08-476-537-84
10	38	50.7	251	1	US-08-485-607-84
11	38	50.7	251	2	US-08-475-879-84
12	38	50.7	251	4	US-09-433-043B-84
13	38	50.7	460	1	US-08-476-008-50
14	38	50.7	460	1	US-08-306-063-50
15	38	50.7	460	1	US-08-833-485-50
16	38	50.7	460	4	US-09-137-440-50
17	38	50.7	621	3	US-08-604-789B-4
18	38	50.7	621	4	US-09-312-721A-4
19	38	50.7	667	4	US-09-094-557-1
20	38	50.7	703	4	US-09-433-043B-124
21	38	50.7	943	4	US-08-808-982-7
22	38	50.7	943	4	US-09-306-902A-7
23	37	49.3	423	2	US-08-955-713-2
24	36	48.0	96	4	US-08-679-493A-147
25	36	48.0	569	1	US-08-306-231-3
26	36	48.0	676	2	US-08-398-590A-40
27	36	48.0	676	4	US-08-894-997-40

28	36	48.0	896	2	US-08-640-389A-10	Sequence 10, Appl
29	36	48.0	896	4	US-08-618-957A-10	Sequence 10, Appl
30	36	48.0	896	2	US-08-693-697-36	Sequence 36, Appl
31	36	48.0	898	4	US-08-588-188-3	Sequence 3, Appl
32	36	48.0	906	2	US-08-640-389A-9	Sequence 9, Appl
33	36	48.0	906	4	US-08-618-957A-9	Sequence 9, Appl
34	36	48.0	908	2	US-08-693-697-33	Sequence 33, Appl
35	36	48.0	908	2	US-08-588-526-3	Sequence 3, Appl
36	36	48.0	958	2	US-08-640-389A-8	Sequence 8, Appl
37	36	48.0	958	4	US-08-618-957A-8	Sequence 8, Appl
38	36	48.0	960	1	US-08-355-886A-8	Sequence 8, Appl
39	36	48.0	960	2	US-08-588-190-3	Sequence 3, Appl
40	36	48.0	960	2	US-08-693-697-8	Sequence 8, Appl
41	36	48.0	960	2	US-08-640-389A-3	Sequence 3, Appl
42	36	48.0	960	3	US-08-693-696-8	Sequence 3, Appl
43	36	48.0	960	4	US-08-618-957A-3	Sequence 3, Appl
44	36	48.0	1165	2	US-08-640-389A-11	Sequence 11, Appl
45	36	48.0	1165	2	US-08-599-455B-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-09-288-143-205  
; Sequence 205, Application US/09288143  
; Patent No. 6433139  
; GENERAL INFORMATION:  
; APPLICANT: Brewer et al.  
; TITLE OF INVENTION: 53 Human Secreted Proteins  
; FILE REFERENCE: P2018P1  
; CURRENT APPLICATION NUMBER: US/09/288,143  
; EARLIER FILING DATE: 1999-04-08  
; EARLIER APPLICATION NUMBER: PCT/US98/21142  
; EARLIER FILING DATE: 1998-10-08  
; EARLIER APPLICATION NUMBER: 60/061,463  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,529  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/071,498  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,527  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,536  
; EARLIER FILING DATE: 1997-10-09  
; EARLIER APPLICATION NUMBER: 60/061,532  
; EARLIER FILING DATE: 1997-10-09  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 205  
; LENGTH: 75  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-288-143-205

Query Match 58.7%; Score 44; DB 4; Length 75;  
Best Local Similarity 66.7%; Pred. No. 3.7;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 2 SSILCSPLPSLS 13  
: | ||||| :  
US-09-177-249-208  
; Sequence 208, Application US/09177249  
; Patent No. 6229064  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Robert L.  
; APPLICANT: Ohad, Nir  
; APPLICANT: Kiyosue, Tomohiro  
; APPLICANT: Yadgar, Ramon

APPLICANT: Margossian, Linda  
APPLICANT: Harada, John  
APPLICANT: Goldberg, Robert B.  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit  
TITLE OF INVENTION: Development In Plants  
FILE REFERENCE: 023070-086120US  
CURRENT APPLICATION NUMBER: US/09/177,249  
EARLIER FILING DATE: 1998-10-22  
EARLIER APPLICATION NUMBER: US 09/071,838  
EARLIER FILING DATE: 1998-05-01  
NUMBER OF SEQ ID NOS: 324  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 208  
LENGTH: 106  
TYPE: PRT  
ORGANISM: Arabidopsis sp.  
US-09-177-249-208

Query Match 57.3%; Score 43; DB 4; Length 106;  
Best Local Similarity 72.7%; Pred. No. 7.7;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 ILCSPLPSLSL 14  
:|||||  
Db 61 MLCSPSPSLHL 71

RESULT 3  
US-09-022-983-2  
Sequence 2, Application US/09022983  
Patent No. 6159731

GENERAL INFORMATION:  
APPLICANT: Yang, Xiaolu  
APPLICANT: Khosravi-Far, Roya  
APPLICANT: Chang, Howard Y.  
APPLICANT: Baltimore, David  
TITLE OF INVENTION: DAXX, A NOVEL FAS-BINDING  
TITLE OF INVENTION: PROTEIN THAT ACTIVATES JNK AND APOPTOSIS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.  
STREET: 600 Atlantic Avenue  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02210

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,983  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/037,919  
FILING DATE: 12-FEB-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/051,753  
FILING DATE: 26-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Van Amsterdam, John R.  
REGISTRATION NUMBER: 40,212  
REFERENCE/DOCKET NUMBER: M0656/7036  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 739 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-09-022-983-2

Query Match 56.0%; Score 42; DB 4; Length 739;  
Best Local Similarity 69.2%; Pred. No. 90;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 SSILCSPLPSLSL 14  
:|||||  
Db 694 TSLCSPSPSLHL 706

RESULT 4  
US-09-328-869-2  
Sequence 2, Application US/09328869B  
Patent No. 6168933

GENERAL INFORMATION:  
APPLICANT: Kaser, Matthew, R.  
APPLICANT: Hillman, Jennifer, L.  
APPLICANT: Baughn, Mariah, R.  
TITLE OF INVENTION: PHOSPHOLIPID TRANSFER PROTEIN  
FILE REFERENCE: PC-0003 US  
CURRENT APPLICATION NUMBER: US/09/328,869B  
CURRENT FILING DATE: 1999-06-08  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PERL Program  
SEQ ID NO 2  
LENGTH: 291  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:

OTHER INFORMATION: 1772859CD1  
US-09-328-869-2

Query Match 52.0%; Score 39; DB 4; Length 291;  
Best Local Similarity 88.9%; Pred. No. 95;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 SPLPSLSLS 15  
:|||||  
Db 242 SPLPSLSLS 250

RESULT 5  
US-09-629-774A-2  
Sequence 2, Application US/09629774A  
Patent No. 6287786

GENERAL INFORMATION:  
APPLICANT: Kaser, Matthew R.  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Baughn, Mariah R.  
TITLE OF INVENTION: PHOSPHOLIPID TRANSFER PROTEIN  
FILE REFERENCE: PC-0003-1 CIP  
CURRENT APPLICATION NUMBER: US/09/629,774A  
CURRENT FILING DATE: 2000-07-31  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PERL Program  
SEQ ID NO 2  
LENGTH: 291  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. 6287786 1772859CD1  
US-09-629-774A-2

Query Match 52.0%; Score 39; DB 4; Length 291;  
Best Local Similarity 88.9%; Pred. No. 95;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 7 SPLPSLSL 15  
DB 242 SPLPSLSL 250

## RESULT 6

US-08-299-187-13  
Sequence 13, Application US/08299187  
Patent No. 5736325  
GENERAL INFORMATION:  
APPLICANT: Manowitz, Paul  
APPLICANT: Poretz, Ronald D.  
APPLICANT: Ricketts, Michael H.  
TITLE OF INVENTION: MARKER FOR INDIVIDUALS SUSCEPTIBLE TO  
TITLE OF INVENTION: ALCOHOLISM  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/299,187  
FILING DATE: 31-AUG-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 601-1-028  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
INDIVIDUAL ISOLATE: peptide from mutant arylsulfatase A  
US-08-299-187-13

Query Match 50.7% Score 38; DB 1; Length 13;  
Best Local Similarity 66.7%; Pred. No. 4.6;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 CSPLPSLSL 14  
DB 1 CAPLPSVTL 9

## RESULT 7

PCT-US95-11114-13  
Sequence 13, Application PC/TUS9511114  
GENERAL INFORMATION:  
APPLICANT: Manowitz, Paul  
APPLICANT: Poretz, Ronald D.  
APPLICANT: Park, David  
APPLICANT: Ricketts, Michael H.

TITLE OF INVENTION: MARKER FOR INDIVIDUALS SUSCEPTIBLE TO  
TITLE OF INVENTION: ALCOHOLISM  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/11114  
FILING DATE: 30-AUG-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/  
FILING DATE: 21-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/299,187  
FILING DATE: 31-AUG-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1158-1-001PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
INDIVIDUAL ISOLATE: peptide from mutant arylsulfatase A  
PCT-US95-11114-13

Query Match 50.7% Score 38; DB 5; Length 13;  
Best Local Similarity 66.7%; Pred. No. 4.6;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 6 CSPLPSLSL 14  
DB 1 CAPLPSVTL 9

## RESULT 8

US-07-956-700B-84  
Sequence 84, Application US/07956700B  
Patent No. 5539092  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5539092th Clark Street  
CITY: Chicago  
STATE: Illinois

COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/956,700B  
FILING DATE: 19921002  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 539092thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-07-956-700B-84

Query Match 50.7%; Score 38; DB 1; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLPSLSLV 16  
DB 181 TSSVLRSPMPGVAV 196

RESULT 9  
US-08-476-537-84  
Sequence 84, Application US/08476537.  
Patent No. 5756290  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5756290th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,537  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5756290thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:

LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-476-537-84

Query Match 50.7%; Score 38; DB 1; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLPSLSLV 16  
DB 181 TSSVLRSPMPGVAV 196

RESULT 10  
US-08-485-607-84  
Sequence 84, Application US/08485607  
Patent No. 5792627  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5792627th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,607  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5792627thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-485-607-84

Query Match 50.7%; Score 38; DB 1; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLPSLSLV 16  
DB 181 TSSVLRSPMPGVAV 196

RESULT 11  
US-08-475-879-84  
Sequence 84, Application US/08475879.  
Patent No. 5972644

Patent No. 5972644 5786170  
GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cynobacterial and Plant Acetyl-CoA  
NUMBER OF SEQUENCES: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: 321 No. 5972644 5786170th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,879  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5972644 5786170chrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ. ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-08-475-879-84

Query Match 50.7%; Score 38; DB 2; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
DB 181 TSSVLKSPMPGVAV 196

RESULT 12  
US-09-433-043B-84  
Sequence 84, Application US/09433043B  
Patent No. 6399342  
GENERAL INFORMATION:  
APPLICANT: HASELKORN, ROBERT  
APPLICANT: GORNICKI, PIOTR  
TITLE OF INVENTION: CYANOBACTERIAL AND PLANT ACETYL-COA CARBOXYLASE  
FILE REFERENCE: ARCD:338US  
CURRENT APPLICATION NUMBER: US/09/433,043B  
CURRENT FILING DATE: 1999-10-25  
PRIOR APPLICATION NUMBER: 08/475,879  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: 07/956,700  
PRIOR FILING DATE: 1992-10-02  
NUMBER OF SEQ ID NOS: 128  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 84  
LENGTH: 251  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Peptide  
US-09-433-043B-84

Query Match 50.7%; Score 38; DB 4; Length 251;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
DB 181 TSSVLKSPMPGVAV 196

RESULT 13  
US-08-476-008-50  
Sequence 50, Application US/08476008  
Patent No. 5627061  
GENERAL INFORMATION:  
APPLICANT: Barry, Gerard F.  
APPLICANT: Kishore, Ganesh M.  
APPLICANT: Padgett, Stephen R.  
APPLICANT: Stallings, William C.  
TITLE OF INVENTION: Glyphosate Tolerant  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
STREET: 700 Chesterfield Village Parkway  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,008  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/306,063  
FILING DATE: 13-SEP-1994  
APPLICATION NUMBER: US 07/749,611  
FILING DATE: 28-AUG-1991  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/576,537  
FILING DATE: 31-AUG-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Hoerner Jr., Dennis R.  
REGISTRATION NUMBER: 30,914  
REFERENCE/DOCKET NUMBER: 38-21(10660)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-6099  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ. ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 460 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-476-008-50

Query Match 50.7%; Score 38; DB 1; Length 460;  
Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 15  
DB 14 SSVNICAPPSKSS 28

RESULT 14  
US-08-306-063-50  
; Sequence 50, Application US/08306063  
; Patent No. 5633435  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyposate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
; STREET: 700 Chesterfield Village Parkway  
; City: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,063  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/576,537  
; FILING DATE: 31-AUG-1990  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoerner Jr., Dennis R.  
; REGISTRATION NUMBER: 30,914  
; REFERENCE/DOCKET NUMBER: 38-21(10660)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)537-6099  
; TELEFAX: (314)537-6047  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 460 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-306-063-50

Query Match 50.7%; Score 38; DB 1; Length 460;  
Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLS 15  
|:::|:| |::|  
Db 14 SSNVICAPGSKSIS 28

RESULT 15  
US-08-833-485-50  
; Sequence 50, Application US/08833485  
; Patent No. 5804425  
; GENERAL INFORMATION:  
; APPLICANT: Barry, Gerard F.  
; APPLICANT: Kishore, Ganesh M.  
; APPLICANT: Padgett, Stephen R.  
; APPLICANT: Stallings, William C.  
; TITLE OF INVENTION: Glyposate Tolerant  
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases  
; NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF  
; STREET: 700 Chesterfield Village Parkway  
; City: St. Louis  
; STATE: Missouri  
; COUNTRY: USA  
; ZIP: 63198  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833,485  
; FILING DATE: 07-APR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/306,063  
; FILING DATE: 13-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,611  
; FILING DATE: 28-AUG-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/576,537  
; FILING DATE: 31-AUG-1990  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hoerner Jr., Dennis R.  
; REGISTRATION NUMBER: 30,914  
; REFERENCE/DOCKET NUMBER: 38-21(15117)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)737-6099  
; TELEFAX: (314)737-6047  
; INFORMATION FOR SEQ ID NO: 50:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 460 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-833-485-50

Query Match 50.7%; Score 38; DB 1; Length 460;  
Best Local Similarity 46.7%; Pred. No. 2.2e+02;  
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLS 15  
|:::|:| |::|  
Db 14 SSNVICAPGSKSIS 28

Search completed: June 25, 2003, 17:18:07  
Job time: 1.85906 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:15:40 ; Search time 1.32438 Seconds  
(without alignments)  
1307.255 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75

Sequence: 1 SSSILCSPLPSLSLV 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCT05\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	16	9 US-10-075-846-74	Sequence 74, Appl
2	75	100.0	312	9 US-10-075-846-15	Sequence 15, Appl
3	75	100.0	431	9 US-10-075-846-4	Sequence 4, Appl
4	47	62.7	83	9 US-09-764-891-3412	Sequence 3412, Ap
5	44	58.7	75	9 US-10-150-111-205	Sequence 205, App
6	44	58.7	508	10 US-09-771-161A-102	Sequence 102, App
7	44	58.7	756	10 US-09-771-161A-101	Sequence 101, App
8	44	58.7	887	10 US-09-771-161A-192	Sequence 192, App
9	44	57.3	106	9 US-09-771-161A-193	Sequence 193, App
10	43	57.3	106	9 US-10-213-512-208	Sequence 208, App
11	43	57.3	106	9 US-09-071-838-208	Sequence 208, App
12	43	57.3	3594	9 US-10-150-821-4	Sequence 4, Appl
13	43	57.3	3594	9 US-09-911-842-4	Sequence 4, Appl
14	41	54.7	216	10 US-09-747-155-49	Sequence 49, Appl
15	41	54.7	310	9 US-10-001-857-196	Sequence 196, App
16	40	53.3	141	10 US-09-746-284-23	Sequence 23, Appl
17	40	53.3	265	10 US-09-764-864-1006	Sequence 1006, Ap
18	40	53.3	400	10 US-09-966-871-86	Sequence 86, Appl
19	40	53.3	400	12 US-10-039-645-86	Sequence 86, Appl

20	39	52.0	68	9 US-10-106-698-5490	Sequence 5490, Ap
21	39	52.0	123	9 US-09-764-891-2743	Sequence 2743, Ap
22	39	52.0	178	10 US-09-811-284-256	Sequence 256, App
23	39	52.0	218	9 US-09-764-891-3395	Sequence 3395, Ap
24	39	52.0	270	10 US-09-841-132-404	Sequence 404, App
25	39	52.0	347	10 US-09-925-301-985	Sequence 985, App
26	39	52.0	359	9 US-10-106-698-4697	Sequence 4697, Ap
27	39	52.0	359	9 US-10-205-823-60	Sequence 60, App
28	38	50.7	25	10 US-09-864-761-36023	Sequence 36023, A
29	38	50.7	41	10 US-09-864-761-47002	Sequence 47002, A
30	38	50.7	76	9 US-10-106-698-5785	Sequence 5785, Ap
31	38	50.7	105	10 US-09-939-907-157	Sequence 157, App
32	38	50.7	141	10 US-09-864-761-37903	Sequence 37903, A
33	38	50.7	346	9 US-10-085-108-22	Sequence 22, Appl
34	38	50.7	362	9 US-09-796-753-142	Sequence 142, App
35	38	50.7	460	9 US-09-464-099A-50	Sequence 50, Appl
36	38	50.7	460	10 US-09-861-666-50	Sequence 50, Appl
37	38	50.7	491	9 US-09-965-529-12	Sequence 12, Appl
38	38	50.7	621	10 US-09-733-300-4	Sequence 4, Appl
39	38	50.7	667	10 US-09-816-127-1	Sequence 1, Appl
40	38	50.7	827	9 US-10-101-464A-915	Sequence 915, App
41	38	50.7	897	9 US-10-138-927-45	Sequence 45, Appl
42	38	50.7	943	9 US-09-933-261-7	Sequence 7, Appl
43	38	50.7	943	9 US-10-256-702-7	Sequence 7, Appl
44	37.5	50.0	126	10 US-09-740-668A-42	Sequence 42, Appl
45	37.5	50.0	300	9 US-10-174-590-548	Sequence 548, App

## ALIGNMENTS

RESULT 1  
US-10-075-846-74  
; Sequence 74, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075, 846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269, 535  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 74  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-075-846-74

Query Match  
Best Local Similarity 100.0%; Score 75; DB 9; Length 16;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 SSSILCSPLPSLSLV 16  
DB 1 SSSILCSPLPSLSLV 16

RESULT 2  
US-10-075-846-15  
; Sequence 15, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075, 846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269, 535

;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 81  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 15  
;; LENGTH: 312  
;; TYPE: PRT  
;; ORGANISM: homo sapiens  
US-10-075-846-15

Query Match 100.0%; Score 75; DB 9; Length 312;  
Best Local Similarity 100.0%; Pred. No. 0.00061;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
|||||

Db 149 SSSIICSPPLSLSLV 164

RESULT 3  
US-10-075-846-4

;; Sequence 4, Application US/10075846  
;; Publication No. US20030032608A1  
;; GENERAL INFORMATION:

;; APPLICANT: Bristol-Myers Squibb Company

;; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT B

;; FILE REFERENCE: D0079 NP IN THE GASTROINTESTINAL TRACT, HGR4, and SPLICE VARIANT THEREOF

;; CURRENT APPLICATION NUMBER: US/10/075,846

;; PRIOR FILING DATE: 2002-02-13

;; PRIOR APPLICATION NUMBER: US 60/269,535

;; PRIOR FILING DATE: 2001-02-16

;; NUMBER OF SEQ ID NOS: 81

;; SOFTWARE: PatentIn version 3.0

;; SEQ ID NO 4

;; LENGTH: 431

;; TYPE: PRT

;; ORGANISM: homo sapiens  
US-10-075-846-4

Query Match 100.0%; Score 75; DB 9; Length 431;  
Best Local Similarity 100.0%; Pred. No. 0.00084;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 16  
|||||

Db 192 SSSIICSPPLSLSLV 207

RESULT 4  
US-09-764-891-3412

;; Sequence 3412, Application US/09764891  
;; Publication No. US20030077808A1  
;; GENERAL INFORMATION:

;; APPLICANT: Rosen et al.

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

;; FILE REFERENCE: PC006

;; CURRENT APPLICATION NUMBER: US/09/764,891

;; PRIOR FILING DATE: 2001-01-17

;; PRIOR APPLICATION data removed - consult PALM or file wrapper

;; NUMBER OF SEQ ID NOS: 10231

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 3412

;; LENGTH: 83

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: SITE

;; LOCATION: (51)

;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-3412

Query Match 62.7%; Score 47; DB 9; Length 83;  
Best Local Similarity 73.3%; Pred. No. 3;

Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSIICSPPLSLSLV 15  
|||||

Db 66 SSSCLVPLPLFLA 80

RESULT 5  
US-10-150-111-205

;; Sequence 205, Application US/10150111  
;; Publication No. US2003007836A1  
;; GENERAL INFORMATION:

;; APPLICANT: Rubin et al.

;; TITLE OF INVENTION: Secreted Protein HPEAD48

;; FILE REFERENCE: P2018P1D1

;; CURRENT APPLICATION NUMBER: US/10/150,111

;; PRIOR FILING DATE: 2002-05-20

;; PRIOR APPLICATION NUMBER: 09/288,143

;; PRIOR FILING DATE: 1999-04-08

;; PRIOR APPLICATION NUMBER: PCT/US98/21142

;; PRIOR FILING DATE: 1998-10-08

;; PRIOR APPLICATION NUMBER: 60/061,463

;; PRIOR FILING DATE: 1997-10-09

;; PRIOR APPLICATION NUMBER: 60/061,529

;; PRIOR FILING DATE: 1997-10-09

;; PRIOR APPLICATION NUMBER: 60/071,498

;; PRIOR FILING DATE: 1997-10-09

;; PRIOR APPLICATION NUMBER: 60/061,527

;; PRIOR FILING DATE: 1997-10-09

;; PRIOR APPLICATION NUMBER: 60/061,536

;; PRIOR FILING DATE: 1997-10-09

;; PRIOR APPLICATION NUMBER: 60/061,532

;; PRIOR FILING DATE: 1997-10-09

;; NUMBER OF SEQ ID NOS: 219

;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 205

;; LENGTH: 75

;; TYPE: PRT

;; ORGANISM: Homo sapiens  
US-10-150-111-205

Query Match 58.7%; Score 44; DB 9; Length 75;  
Best Local Similarity 66.7%; Pred. No. 7.9;

Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSIICSPPLSLSLV 13  
|||||

Db 15 TSQCLSPMPSLN 26

RESULT 6  
US-09-771-161A-102

;; Sequence 102, Application US/09771161A  
;; Patent No. US2002010811A1  
;; GENERAL INFORMATION:

;; APPLICANT: LEVINE, et al.

;; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

;; FILE REFERENCE: 802620-2005.1

;; CURRENT APPLICATION NUMBER: US/09/771,161A

;; PRIOR FILING DATE: 2001-01-26

;; PRIOR APPLICATION NUMBER: 09/724,676

;; PRIOR FILING DATE: 2000-11-28

;; PRIOR APPLICATION NUMBER: 136776

;; PRIOR FILING DATE: 2000-06-15

;; PRIOR APPLICATION NUMBER: 135619

;; PRIOR FILING DATE: 2000-04-12

;; NUMBER OF SEQ ID NOS: 273

;; SOFTWARE: PatentIn version 3.0

;; SEQ ID NO 102

;; LENGTH: 508

;; TYPE: PRT

;; ORGANISM: Homo sapiens  
US-09-771-161A-102



Query Match 58.7%: Score 44; DB 10; Length 508;  
Best Local Similarity 69.2%: Pred. No. 54;  
Matches 9: Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1  
DB 442 SSOITSPPLSPS 454

## RESULT 7

US-09-771-161A-101  
; Sequence 101, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771.161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 101  
; LENGTH: 756  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-101

Query Match 58.7%: Score 44; DB 10; Length 756;  
Best Local Similarity 69.2%: Pred. No. 80;  
Matches 9: Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1  
DB 442 SSOITSPPLSPS 454

## RESULT 8

US-09-771-161A-192  
; Sequence 192, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771.161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 192  
; LENGTH: 887  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-192

Query Match 58.7%: Score 44; DB 10; Length 887;  
Best Local Similarity 69.2%: Pred. No. 94;  
Matches 9: Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1

DB 442 SSOITSPPLSPS 454

## RESULT 9

US-09-771-161A-193  
; Sequence 193, Application US/09771161A  
; Patent No. US20020110811A1  
; GENERAL INFORMATION:  
; APPLICANT: LEVINE, et al.  
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES  
; FILE REFERENCE: 802620-2005.1  
; CURRENT APPLICATION NUMBER: US/09/771.161A  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/724,676  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 136776  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 135619  
; PRIOR FILING DATE: 2000-04-12  
; NUMBER OF SEQ ID NOS: 273  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 193  
; LENGTH: 887  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-771-161A-193

Query Match 58.7%: Score 44; DB 10; Length 887;  
Best Local Similarity 69.2%: Pred. No. 94;  
Matches 9: Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSICSPPLSPS 13  
111:11111:1  
DB 442 SSOITSPPLSPS 454

## RESULT 10

US-10-213-512-208  
; Sequence 208, Application US/10213512  
; Publication No. US20030110536A1  
; GENERAL INFORMATION:  
; APPLICANT: Fischer, Robert L.  
; APPLICANT: Ohad, Nir  
; APPLICANT: Kiyosue, Tomohiro  
; APPLICANT: Yadegari, Ramin  
; APPLICANT: Matgossian, Linda  
; APPLICANT: Harada, John  
; APPLICANT: Goldberg, Robert B.  
; TITLE OF INVENTION: The Regents of the University of California  
; TITLE OF INVENTION: Combinations of Nucleic Acids That Control Seed and  
; FILE REFERENCE: 023070-086100S  
; CURRENT APPLICATION NUMBER: US/10/213.512  
; PRIOR FILING DATE: 2002-08-06  
; PRIOR APPLICATION NUMBER: US/09/177.206  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: US 09/071.838  
; PRIOR FILING DATE: 1998-05-01  
; NUMBER OF SEQ ID NOS: 324  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 208  
; LENGTH: 106  
; TYPE: PRT  
; ORGANISM: Arabidopsis sp.  
US-10-213-512-208

Query Match 57.3%: Score 43; DB 9; Length 106;  
Best Local Similarity 72.7%: Pred. No. 16;  
Matches 8: Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 TLCSPLPSLSL 14  
:11111111  
DB 61 MLCSPPLSLHL 71

RESULT 11  
US-09-071-838-208  
Sequence 208, Application US/09071838  
Patent No. US20020152501A1  
GENERAL INFORMATION:  
APPLICANT: Fischer, Robert L.  
APPLICANT: Ohad, Nir  
APPLICANT: Kiyosue, Tomohiro  
APPLICANT: Yadegari, Ramtin  
APPLICANT: Margossian, Linda  
APPLICANT: Harada, John  
APPLICANT: Goldberg, Robert B.  
TITLE OF INVENTION: Nucleic Acids That Control Seed and  
TITLE OF INVENTION: Fruit Development in Plants  
NUMBER OF SEQUENCES: 324  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/071,838  
FILING DATE: 01-MAY-1998  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bastian, Kevin L.  
REGISTRATION NUMBER: 34,774  
REFERENCE/DOCKET NUMBER: 023070-08610005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 208:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 106 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-071-838-208

Query Match 57.3%; Score 43; DB 10; Length 106;  
Best Local Similarity 72.7%; Pred. No. 16;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 IICSPPLSL 14  
DB 61 MLCSPPLSL 71

RESULT 12  
US-10-150-821-4  
Sequence 4, Application US/10150821  
Patent No. US20020192758A1  
GENERAL INFORMATION:  
APPLICANT: Weichert, Andrew A.  
APPLICANT: Elliott, Gary S.  
TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: 01017/37592  
CURRENT APPLICATION NUMBER: US/10/150,821  
CURRENT FILING DATE: 2002-05-16  
PRIOR APPLICATION NUMBER: US/09/911,842  
PRIOR FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: US 60/222,438  
PRIOR FILING DATE: 2000-08-01

NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 3594  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1757)..()  
OTHER INFORMATION: Xaa = any or unknown amino acid  
US-10-150-821-4

Query Match 57.3%; Score 43; DB 9; Length 3594;  
Best Local Similarity 62.5%; Pred. No. 5.4e+02;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSIICSPPLSLSV 16  
DB 1649 SSMFCSDCPSLGCV 1664

RESULT 13  
US-09-911-842-4  
Sequence 4, Application US/09911842  
Patent No. US20020151483A1  
GENERAL INFORMATION:  
APPLICANT: Weichert, Andrew A.  
APPLICANT: Elliott, Gary S.  
TITLE OF INVENTION: C3B/C4B COMPLEMENT RECEPTOR-LIKE MOLECULES AND USES THEREOF  
FILE REFERENCE: 01017/37592  
CURRENT APPLICATION NUMBER: US/09/911,842  
CURRENT FILING DATE: 2001-07-24  
PRIOR APPLICATION NUMBER: US 60/222,438  
PRIOR FILING DATE: 2000-08-01  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 3594  
TYPE: PRT  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1757)..()  
OTHER INFORMATION: Xaa = any or unknown amino acid  
US-09-911-842-4

Query Match 57.3%; Score 43; DB 10; Length 3594;  
Best Local Similarity 62.5%; Pred. No. 5.4e+02;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SSIICSPPLSLSV 16  
DB 1649 SSMFCSDCPSLGCV 1664

RESULT 14  
US-09-747-155-49  
Sequence 49, Application US/09747155  
Patent No. US20020151692A1  
GENERAL INFORMATION:  
APPLICANT: Rouquier, Sylvie  
APPLICANT: Giorgi, Dominique  
TITLE OF INVENTION: No. US20020151692A1el Polypeptides and Nucleic Acids Encoding  
FILE REFERENCE: 19904-008 (C009B6834US)  
CURRENT APPLICATION NUMBER: US/09/747,155  
CURRENT FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: 60/171,746  
PRIOR FILING DATE: 1999-12-22  
NUMBER OF SEQ ID NOS: 431  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 49  
LENGTH: 216  
TYPE: PRT

Page 5

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:55 ; Search time 1.07383 Seconds  
(without alignments)  
1432.403 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207  
Perfect score: 75  
Sequence: 1 SSSIICSPPLPSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*\*\*  
1: pirl:\*\*\*  
2: pirl:\*\*\*  
3: pirl:\*\*\*  
4: pirl:\*\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	65.3	490	2 T49096	hypothetical prote
2	46	61.3	117	2 C85683	probable IS encode
3	44	58.7	887	1 S57219	1-phosphatidylinos
4	42	56.0	91	1 MNVUAV	nonstructural prot
5	42	56.0	114	2 A72535	hypothetical prote
6	42	56.0	141	2 B82845	hypothetical prote
7	42	56.0	478	2 H71345	probable polynucle
8	42	56.0	1010	2 T41077	hypothetical struc
9	41	54.7	396	2 T02483	probable protein p
10	41	54.7	944	2 T38130	probable helicase
11	41	54.7	2123	2 F86348	hypothetical prote
12	40	53.3	247	2 T22881	hypothetical prote
13	40	53.3	262	2 D71557	hypothetical prote
14	40	53.3	313	2 D96616	hypothetical prote
15	40	53.3	798	2 S09867	hypothetical prote
16	40	53.3	965	2 T38430	hypothetical prote
17	39.5	52.7	227	2 S35240	hypothetical prote
18	39	52.0	261	2 E81520	heat shock protein
19	39	52.0	270	2 C72019	conserved hypotet
20	39	52.0	270	2 B86604	CT764 hypothetical
21	39	52.0	449	2 T16259	hypothetical prote
22	39	52.0	509	2 A36392	segmentation prote
23	39	52.0	647	2 A37086	beta-galactosidase
24	39	52.0	740	2 S55140	hypothetical prote
25	39	52.0	841	1 S24462	probable 3',5'-cyc
26	39	52.0	918	1 D88544	protein R08D7 6 [1
27	38	50.7	100	2 AG0379	acetylactate synth
28	38	50.7	107	2 A72456	NADH dehydrogenase
29	38	50.7	120	2 AC2286	NADH dehydrogenase

30	38	50.7	124	2 A54773	Ac1 protein - mous
31	38	50.7	133	2 A85093	hypothetical prote
32	38	50.7	141	2 F71555	hypothetical prote
33	38	50.7	185	2 T10677	pathogenesis-relat
34	38	50.7	201	2 G68736	hypothetical prote
35	38	50.7	231	2 A82538	hypothetical prote
36	38	50.7	297	2 T27206	hypothetical prote
37	38	50.7	306	2 T49541	hypothetical prote
38	38	50.7	328	2 JN0882	gonadotropin-rela
39	38	50.7	353	2 T36089	probable lacI-fam1
40	38	50.7	449	2 B85069	hypothetical prote
41	38	50.7	491	2 T46915	hypothetical prote
42	38	50.7	554	2 T41612	amino acid permeas
43	38	50.7	576	2 T11046	NADH2 dehydrogenas
44	38	50.7	621	2 S27752	anthranilate synth
45	38	50.7	621	2 J01685	anthranilate synth

ALIGNMENTS

RESULT 1  
T49096  
hypothetical protein F4F15.280 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C:Accession: T49096  
R:Alcaraz, J.P.; Clabault, G.; Cottet, A.; Maché, R.; Mewes, H.W.; Rudd, S.; Lemcke, submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z25015  
A:Accession: T49096  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-490 <ALC>  
A:Cross-references: EMBL:AL049711; GSPDB:GN00061; ATSP:F4F15.280  
C:Genetics:  
A:Experimental source: cultivar Columbia; BAC clone F4F15  
A:Gene: ATSP:F4F15.280  
A:Map position: 3  
A:Insertions: 70/2; 433/3

Query Match 65.3%; Score 49; DB 2; Length 490;  
Best Local Similarity 73.3%; Pred. No. 2.9;  
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 1 SSSIICSPPLPSLS 15  
129 SSSIIMDPVPPLSL 143

RESULT 2  
C85683  
Probable IS encoded protein Z1827 [Imported] - Escherichia coli (strain O157:H7, subs  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: C85683  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: AB5480; MUID:21074935; PMID:11206551  
A:Accession: C85683  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-117 <STO>  
A:Cross-references: GB:AE005174; NID:G12514746; PIDN:AAG55927.1; GSPDB:GN00145; UWGP:  
C:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z1827

Query Match 61.3%; Score 46; DB 2; Length 117;  
Best Local Similarity 53.3%; Pred. No. 2.2;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSILCSPPLSLSV 16  
:||:||||:|  
Db 12 TSILCSPMTSLKTSI 26

## RESULT 3

S57219  
1-phosphatidylinositol 3-kinase (EC 2.7.1.137) Vps34-type [validated] - human  
C:Species: Homo sapiens (man)  
C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 01-Feb-2002  
C:Accession: S57219  
R:Volinia, S.; Dhond, R.; Vanhaesebroeck, B.; MacDougall, L.K.; Stein, R.; Zvelebil, M.J.  
EMBO J. 14, 3339-3348, 1995  
A:Title: A human phosphatidylinositol 3-kinase complex related to the yeast Vps34p-Vps15  
A:Reference number: S57219; MUID:95354652; PMID:7628435  
A:Accession: S57219  
A:Molecule type: mRNA  
A:Residues: 1-887 <VOL>  
A:Function:  
A:Description: converts ATP and 1-phosphatidyl-1D-myo-inositol to ADP and 1-phosphatidyl-  
A:Pathway: required for vacuolar sorting and segregation; involved in both internalizati  
A>Note: specific for phosphatidylinositol, inactive on phosphatidylinositol-3-phosphate  
C:Keywords: phosphotransferase; protein trafficking

Query Match 58.7%; Score 44; DB 1; Length 887;  
Best Local Similarity 69.2%; Pred. No. 34;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SSILCSPPLSLSV 13  
||:||||:|  
Db 442 SSQITSPPLSPSVS 454

## RESULT 4

MANVAV  
nonstructural protein NS-S - alno virus  
C:Species: alno virus  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C:Accession: S07943  
R:AKashl, H.; Gay, M.; Ihara, T.; Bishop, D.H.L.  
Virus Res. 1, 51-63, 1984  
A:Title: Localized conserved regions of the S RNA gene products of bunyaviruses are reve  
A:Reference number: S07414; MUID:85170502; PMID:6532000  
A:Accession: S07943  
A:Molecule type: genomic RNA  
A:Residues: 1-91 <AKA>  
A:Cross-references: EMBL:M22011; NID:9210064; PIDN:AAA42544.1; PID:9210066  
C:Genetics:  
A:Map position: segment 5  
C:Superfamily: bunyavirus nonstructural protein  
C:Keywords: nonstructural protein

Query Match 56.0%; Score 42; DB 1; Length 91;  
Best Local Similarity 57.1%; Pred. No. 7.6;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 SSILCSPPLSLSV 14  
:||||:|  
Db 73 TQILCQTPSLST 86

## RESULT 5

A72535  
hypothetical protein APE1566 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: A72535  
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: A72535  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-114 <KAM>  
A:Cross-references: DDBJ:AP000062; NID:95105244; PIDN:BAA80566.1; PID:dl044352; PID:9  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1566  
C:Superfamily: Aeropyrum pernix hypothetical protein APE1566

Query Match 56.0%; Score 42; DB 2; Length 114;  
Best Local Similarity 54.5%; Pred. No. 9.5;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 6 CSPPLSLSLSV 16  
||:||||:|  
Db 28 CTPPSMNLST 38

## RESULT 6

B82845  
hypothetical protein XF0131 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82845  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82845  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-141 <STM>

A:Cross-references: GB:AE003866; GB:AE003849; NID:99104906; PIDN:AAF82944.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Brienes, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer  
as-Neto, E.; Docena, C.; El-Dorriy, H.; Fachinani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr  
J.D.; Junqueira, M.L.; Kempner, E.L.; Kitejima, J.P.; Klieger, J.E.; Kuramae, E.E.; La  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0131

Query Match 56.0%; Score 42; DB 2; Length 141;  
Best Local Similarity 57.1%; Pred. No. 12;  
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSILCSPPLSLSV 14  
:||||:|  
Db 9 TSLRCAPPLSLST 22

## RESULT 7

H71345  
probable polynucleotide adenylyltransferase (pcna) - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
C:Accession: H71345  
R:Riferson, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; M  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998

A>Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A:Reference number: A71250; MUID:9832770; PMID:9665876  
A:Accession: H71345  
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-478 <COL>  
A:Cross-references: GB:AE001208; GB:AE000520; NID:g3322538; PIDN:AA05258.1; PID:g332254  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0270

Query Match 56.0%; Score 42; DB 2; Length 478;  
Best Local Similarity 53.8%; Pred. No. 39;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

OY 2 SSILCSPLPSLSL 14  
DB 203 SKMLCTPRPSIAL 215

## RESULT 8

T41077

hypothetical structural protein - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T41077  
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Zimmermann, W.; Wambutt, R.  
submitted to the EMBL Data Library, August 1999  
A:Reference number: Z21822

A:Accession: T41077  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1010 <MCDD>

A:Cross-references: EMBL:AL109957; PIDN:CAB53076.1; GSPDB:GN00068; SPDB:SPCC16A11.04  
A:Experimental source: strain 972h; cosmid c16A11  
C:Genetics:  
A:Gene: SPDB:SPCC16A11.04  
A:Map position: 3

Query Match 56.0%; Score 42; DB 2; Length 1010;  
Best Local Similarity 62.5%; Pred. No. 81;  
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSILCSPLPSLSLSV 16  
DB 572 SKINSPPSPISQISLV 587

## RESULT 9

T02483

probable protein phosphatase 2C At2g30020 [imported] - *Arabidopsis thaliana*  
N:Alternate names: hypothetical protein F23F1.6  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 03-Jun-2002  
C:Accession: T02483; E84703  
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul  
submitted to the EMBL Data Library, August 1998  
A:Description: The *Arabidopsis thaliana* chromosome II BAC F23F1 genomic sequence.  
A:Reference number: Z14675

A:Accession: T02483  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-396 <ROU>

A:Cross-references: EMBL:AC004680; NID:g3420043; PID:g3420049  
A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Motil, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: E84703  
A:Status: Preliminary

A:Molecule type: DNA  
A:Residues: 1-396 <STO>  
A:Cross-references: GB:AE002093; NID:g3420049; PIDN:AA01850.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g30020; F23F1.6  
A:Map position: 2  
A:Introns: 168/3  
C:Superfamily: *Arabidopsis thaliana* hypothetical protein F7A7.220

Query Match 54.7%; Score 41; DB 2; Length 396;  
Best Local Similarity 66.7%; Pred. No. 47;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSILCSPLPSLSL 16  
DB 24 SSILSPQESLSLTL 38

## RESULT 10

T38130

probable helicase - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T38130  
R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1997  
A:Reference number: Z21772

A:Accession: T38130  
A:Status: Preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-944 <BAD>

A:Cross-references: EMBL:Z95334; PIDN:CAB08602.1; GSPDB:GN00066; SPDB:SPAC20G8.08C  
A:Experimental source: strain 972h; cosmid c20G8  
C:Genetics:  
A:Gene: SPDB:SPAC20G8.08C  
A:Map position: 1

A:Introns: 538/3; 583/3; 717/3; 894/3

Query Match 54.7%; Score 41; DB 2; Length 944;  
Best Local Similarity 60.0%; Pred. No. 11e+02;  
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSILCSPLPSLSLS 15  
DB 32 SSIFCSPIQQLDLS 46

## RESULT 11

F86348

hypothetical protein F24J8.17 [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C:Accession: F86348  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Matzla  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A>Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: AB6141; MUID:21016719; PMID:11130712

A:Accession: F86348  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-2123 <STO>

A:Cross-references: GB:AE005172; NID:g9454580; PIDN:AA07903.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1

Query Match 54.7%; Score 41; DB 2; Length 2123;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:59 ; Search time 0.644295 Seconds

(without alignments)  
1029.995 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Sequence: 1 SSSILCSPLPSLSLV 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	56.0	91	VNST_AINOV	P12413 alno virus.
2	40	53.3	401	OPRM_PIG	Q95247 sus scrofa
3	40	53.3	798	HEPA_HCMVA	P16827 human cytom
4	39.5	52.7	227	HS2C_ARATH	P31170 arabidopsi
5	39	52.0	509	RUNT_DROME	P22814 drosophila
6	39	52.0	647	BGAL_MOUSE	P23780 mus musculu
7	39	52.0	739	AABL_MOUSE	Q9w776 mus musculu
8	39	52.0	740	YNO4_YEAST	P53307 saccharomyc
9	39	52.0	918	YNE6_CAEEL	P30645 caenorhabd1
10	38	50.7	120	NU3C_ANASP	Q44239 anabaena sp
11	38	50.7	124	APCA_MOUSE	Q61268 mus musculu
12	38	50.7	328	GRHR_SHEEP	P32237 ovis aries
13	38	50.7	464	IRUS_SHEEP	Q31396 bukhoiderl
14	38	50.7	554	YCT3_SCHPO	Q59813 schizosach
15	38	50.7	621	TRPX_ARATH	P32069 arabidopsi
16	38	50.7	703	PCCA_HUMAN	P05165 homo sapien
17	38	50.7	1586	AROI_EMENT	P07547 e.pentalunc
18	38	50.7	4092	DYHC_YEAST	P36022 saccharomyc
19	37.5	50.0	220	RK21_ARATH	P51412 arabidopsi
20	37	49.3	168	PLAT_POPNI	P00299 populus nig
21	37	49.3	168	PLAT_POPNI	P11970 populus nig
22	37	49.3	176	Y069_TREPA	Q83108 treponema p
23	37	49.3	288	VG27_HSV1	Q00113 ictaluriid h
24	37	49.3	365	Y023_CAEEL	Q09452 caenorhabd1
25	37	49.3	773	GLGB_SYNP	P16954 synecococc
26	37	49.3	803	RIRI_CRYPV	Q61065 cryptospori
27	37	49.3	817	HUNB_MUSDO	Q01778 musca domes
28	37	49.3	1229	KPB2_FUGRU	Q9w6r1 fugu rubrip
29	37	49.3	1770	RII5_YEAST	P43365 saccharomyc
30	36.5	48.7	474	MIG1_KLULA	P50898 klyveromyc
31	36.5	48.7	476	EFNU_ARATH	P17745 arabidopsi
32	36.5	48.7	1609	FIG2_YEAST	P23553 saccharomyc
33	36	48.0	96	GRO_RAT	P14095 rattus norv

34	36	48.0	108	SLIB_HUMAN	P01286 homo sapien
35	36	48.0	264	CERC_SCHMA	P12546 schistosoma
36	36	48.0	271	FRAL_HUMAN	P15407 homo sapien
37	36	48.0	274	PSBS_SPTOL	Q02060 spinacia o1
38	36	48.0	275	FRAL_RAT	P10158 rattus norv
39	36	48.0	430	MS12_AGRRH	P50201 agrobacteri
40	36	48.0	440	FU10_ARATH	Q951p6 arabidopsi
41	36	48.0	475	ETS6_DROME	P29776 drosophila
42	36	48.0	533	FUT5_ARATH	Q951p4 arabidopsi
43	36	48.0	675	NED1_MOUSE	P33215 mus musculu
44	36	48.0	956	CB31_YEAST	P32504 saccharomyc
45	36	48.0	1165	LEPR_HUMAN	P48357 homo sapien

## ALIGNMENTS

RESULT 1  
VNST\_AINOV STANDARD: PRT: 91 AA.  
AC P12413;  
DT 01-OCT-1989 (rel. 12, Created)  
DT 01-OCT-1989 (rel. 12, Last sequence update)  
DT 01-JAN-1990 (rel. 13, Last annotation update)  
DE Nonstructural protein NS-S.  
OS Alno virus.  
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.  
OX NCBI\_TaxID=11582;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85170502; PubMed=6532000;  
RA Akashi H., Gay M., Ihara T., Bishop D.H.L.;  
RT "Localized conserved regions of the S RNA gene products of  
RT bunyaviruses are revealed by sequence analyses of the Simbu serogroup  
RT Alno virus.";  
RL Virus Res. 1:51-63(1984).  
CC -!- SIMILARITY: 35% TO LA CROSSE BUNYAVIRUS NONSTRUCTURAL PROTEIN NS-S  
CC AND TO SNOWSHOE HARE BUNYAVIRUS NS-S.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: M22011; AAA42544.1; -  
CC PIR: S07943; MNVUAV.  
CC InterPro: IPR000797; Bunya\_NSS.  
CC Pfam: PF01104; Bunya\_NS-S; 1.  
CC ProDom: PD002170; Bunya\_NSS; 1.  
CC Nonstructural protein.  
SQ SEQUENCE 91 AA; 10503 MW; 6B9195505744BA0 CRC64;  
  
Query Match 56.0%; Score 42; DB 1; Length 91;  
Best Local Similarity 57.1%; Pred. No. 1.7;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
OY 1 SSSILCSPLPSLSL 14  
DB 73 TQOILCOTLPRLSI 86  
  
RESULT 2  
OPRM\_PIG STANDARD: PRT: 401 AA.  
AC Q95247;  
DT 01-NOV-1997 (rel. 35, Created)  
DT 01-NOV-1997 (rel. 35, Last sequence update)  
DT 15-JUL-1999 (rel. 38, Last annotation update)  
DE Mu-type oploid receptor (MOR-1).  
GN OPRM1.

OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain cortex;  
 RC MEDLINE=99032385; PubMed=9817447;  
 RA Pampusch M.S., Osinski M.A., Brown D.R., Murtaugh M.P.;  
 RT "The porcine mu opioid receptor: molecular cloning and mRNA  
 RT distribution in lymphoid tissues.";  
 RL J. Neuroimmunol. 90:192-198(1998).  
 CC -1- FUNCTION: INHIBITS NEUROTRANSMITTER RELEASE BY REDUCING CALCIUM  
 CC ION CURRENTS AND INCREASING POTASSIUM ION CONDUCTANCE. RECEPTOR  
 CC FOR BETA-ENDORPHIN.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L38645; AAB33770.1; -  
 CC InterPro: IPR000276; GPCR\_Rhodpsn.  
 CC Pfam: PF00001; 7tm.1; 1.  
 DR PRINTS: PR00237; GPCRHOPOPSN.  
 DR PROSITE: PS00237; G-PROTEIN\_RECP\_F1\_1; 1.  
 DR PROSITE: PS00262; G-PROTEIN\_RECP\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.  
 FT DOMAIN 1 67 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 68 97 1 (POTENTIAL).  
 FT DOMAIN 98 106 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 107 124 2 (POTENTIAL).  
 FT DOMAIN 125 146 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 147 166 3 (POTENTIAL).  
 FT DOMAIN 167 196 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 197 212 4 (POTENTIAL).  
 FT DOMAIN 213 237 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 238 260 5 (POTENTIAL).  
 FT DOMAIN 261 283 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 284 306 6 (POTENTIAL).  
 FT DOMAIN 307 314 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 315 331 7 (POTENTIAL).  
 FT DOMAIN 332 401 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 143 220 BY SIMILARITY.  
 FT LIPID 354 354 PALMITATE (POTENTIAL).  
 FT CARBOHYD 9 9 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 12 12 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 401 AA; 45098 MW; 1A208BC5ED7F83FA CRC64;

Query Match 53.3%; Score 40; DB 1; Length 401;  
 Best local Similarity 61.5%; Pred. No. 18;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSIICSPPLPSLS 13  
 Db 18 SPSSMCSVPSPS 30

RESULT 3  
 ID HEPA\_HCMVA STANDARD; PRT; 798 AA.  
 AC P16827;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE DNA helicase/primase complex associated protein.  
 GN UL102.  
 OS Human cytomegalovirus (strain AD169).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Cytomegalovirus.  
 OX NCBI\_TaxID=10360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90269039; PubMed=2161319;  
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
 RA Horsnell T., Hutchinson C.A. III, Kouzarides T., Martignetti J.A.,  
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Bartell B.G.;  
 RT "Analysis of the protein-coding content of the sequence of human  
 RT cytomegalovirus strain AD169.";  
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
 CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.  
 CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,  
 CC EHV-1 54, VZV 52 AND HCMV 102.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X17403; CAA35338.1; -  
 CC PIR: S09867; S09867.  
 DR PIR: S09867; S09867.  
 KW DNA replication.  
 SQ SEQUENCE 798 AA; 85613 MW; 6AA4E14B11E1F5EB CRC64;

Query Match 53.3%; Score 40; DB 1; Length 798;  
 Best local Similarity 60.0%; Pred. No. 39;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSIICSPPLPSLS 15  
 Db 731 SSVSLASPLSVTSS 745

RESULT 4  
 ID HS2C\_ARATH STANDARD; PRT; 227 AA.  
 AC P31170;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Small heat shock protein, chloroplast precursor.  
 GN HSP21 OR AT4G27670 OR T29A15.160.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia; TISSUE=Leaf;  
 RC MEDLINE=91246118; PubMed=2038305;  
 RA Chen Q., Vierling E.;  
 RT "Analysis of conserved domains identifies a unique structural feature  
 RT of a chloroplast heat shock protein.";  
 RL Mol. Gen. Genet. 226:425-431(1991).  
 RN [2]  
 RP ERRATUM.  
 RC MEDLINE=91360086; PubMed=1886617;  
 RA Chen Q., Vierling E.;  
 RL Mol. Gen. Genet. 228:328-328(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RC MEDLINE=93302705; PubMed=8391109;

RA OsterYoung K.W., Sundberg H., Vierling E.;  
 RT "Poly(A) tail length of a heat shock protein RNA is increased by  
 RT severe heat stress, but intron splicing is unaffected."  
 RL Mol. Gen. Genet. 239:323-333(1993).  
 RM (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV, Columbia;  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,  
 RA Poll T., Dueterhoeft A., Stiekema W., Entlin K.D., Terryn N.,  
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Macho R., Mueller M.,  
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohnselt J., Zimmermann W., Wedler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,  
 RA Braeken M., Weltjens I., Voet M., Bastlaens I., Aert R., Defoor E.,  
 RA Weltzenegger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,  
 RA Moolman P., Klein lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Beneliser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Blyssnaert C., Gielen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,  
 RA Petterit A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,  
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Faltmann B., Grandenath K., Dauner D., Herzl A.,  
 RA Neumann S., Argitlout A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chifbons T., Weber N., Vandenbol M., Baques M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,  
 RA Frishman D., Haese D., Lemcke K., Mewes H.-W., Stocker S.,  
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dedha N., Gnoj L., Schutz K., Huang E., Spiegel L.,  
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,  
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
 RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
 RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,  
 RA Du H., All J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
 RA Antonolou B., Zidani M., Strong C., Sun H., Lamar B., Yordan C.,  
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
 RA Grenat S., Shohdy N., Hasegawa A., Hameed A., Lohdi M., Johnson A.,  
 RA Chen E., Marra M., Martienssen R., McCombie M.R.;  
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 402:769-777(1999).  
 CC -1- SUBCELLULAR LOCATION: Chloroplast.  
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)  
 CC FAMILY.  
 CC -1- SIMILARITY: PLANTS HAVE FOUR CLASSES OF HSP20. CLASS I AND II  
 CC WHICH ARE CYTOPLASMIC; CLASS III WHICH IS IN THE CHLOROPLAST  
 CC AND CLASS IV WHICH IS IN THE ENDOPLASMIC RETICULUM. THIS PROTEIN BELONGS  
 CC TO CLASS III.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X54102; CAA38036.1; -  
 CC DR EMBL: M94455; AAA32818.1; -  
 CC DR EMBL: AL161571; CAB81417.1; -

DR EMBL: AL035602; CAB8279.1; -  
 DR PIR: S16527; S16527.  
 DR PIR: S16005; S16005.  
 DR PIR: S35240; S35240.  
 DR InterPro: IPR002068; HSP20.  
 DR Pfam: PF00011; HSP20; 1.  
 DR PROSITE: PS01031; HSP20; 1.  
 KW Heat shock; Chloroplast; Transit peptide.  
 FT TRANSIT 1 19  
 FT CHAIN 20 227  
 FT SEQUENCE 227 AA; 25344 MW; B25400AF6C01972E CRC64;  
 SO  
 Query Match 52.7%; Score 39.5; DB 1; Length 227;  
 Best Local Similarity 62.5%; Pred. No. 12;  
 Matches 10; Conservative 3; Mismatches 2; Indels 1; Gaps 1;  
 Db 8 AASALCSPLAPSPSVS 23  
 QY 1 SSSILCSPL-PSLSLS 15  
 : : | | | | | : : :  
 DB 8 AASALCSPLAPSPSVS 23  
 RESULT 5  
 RUNT\_DROME STANDARD; PRT: 509 AA.  
 AC P22814;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Segmentation protein Runt.  
 GN Runt.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91065517; PubMed=2249771;  
 RA Kania M.A., Bonner A.S., Duffly J.B., Gergen J.P.;  
 RT "The Drosophila segmentation gene runt encodes a novel nuclear  
 RT regulatory protein that is also expressed in the developing nervous  
 RT system."  
 RL Genes Dev. 4:1701-1713(1990).  
 CC -1- SIMILARITY TO AML.  
 CC MEDLINE=92220161; PubMed=1560822;  
 RA Daga A., Tighe J.E., Calabi F.;  
 RT "Leukaemia/Drosophila homology."  
 RL Nature 356:484-484(1992).  
 CC -1- FUNCTION: PLAYS A PIVOTAL ROLE IN REGULATING THE EXPRESSION OF  
 CC OTHER PAIR-RULE GENES SUCH AS EVE, FTZ, AND H.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: DEVELOPING CENTRAL AND PERIPHERAL NERVOUS  
 CC SYSTEM.  
 CC -1- DEVELOPMENTAL STAGE: MOST ABUNDANTLY EXPRESSED AT THE BLASTODERM  
 CC STAGE OF EMBRYOGENESIS.  
 CC -1- SIMILARITY: CONTAINS 1 RUNT DOMAIN.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: X56432; CAA39817.1; -  
 CC DR PIR: A36392; A36392.  
 CC HSSP: 060472; ICMD.  
 CC TRANSPAC: T01066;  
 CC Flybase: FBgn0003300; run.  
 CC InterPro: IPR000040; AML1\_Runt.  
 CC Pfam: PF00853; Runt; 1.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF151517; AAD43327.1; -  
DR MGI: 1337130; Ap4b1.  
DR InterPro: IPR002553; Adaptin\_N.  
DR Pfam: PF01602; Adaptin\_N; 1.  
DR Coated pits; Endocytosis.  
SQ SEQUENCE 739 AA; 82618 MW; 1007D972BF8D0897 CRC64;

Query Match 52.0%; Score 39; DB 1; Length 739;  
Best Local Similarity 58.8%; Pred. No. 53;  
Matches 10; Conservative 1; Mismatches 4; Indels 2; Gaps 1;

OY 1 SSSIICSP--PSLSLS 15  
Db : ||||| ||| ||  
529 TKQITCSKSPSPSLGLS 545

RESULT 8  
YNO4\_YEAST STANDARD: PRT: 740 AA.  
ID YNO4\_YEAST STANDARD: PRT: 740 AA.  
AC P53907;  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE Hypothetical 84.2 kDa protein in MFA2 intergenic region.  
GN YNL144C OR N1205 OR N1790.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C;  
RX MEDLINE=96109932; PubMed=8619318;  
RA Maillet L., Busseteau F., Jacquet M.;  
RT "A 43.5 kb segment of yeast chromosome XIV, which contains MFA2,  
MEP2, CAP/SRY2, NAM9, FKBI/FPRI/RBP1, MOM22 and CPT1, predicts an  
RT adenosine deaminase gene and 14 new open reading frames.";  
RL Yeast 11:1195-1209(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288C / FY1679;  
RX MEDLINE=96287653; PubMed=8686380;  
RA Naar F., Becam A.-M., Herbert C.J.;  
RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals  
RT 24 complete open reading frames: 18 correspond to new genes, one of  
RT which encodes a protein similar to the human myotonic dystrophy  
RT kinase.";  
RL Yeast 12:169-175(1996).  
CC -1- SIMILARITY: TO YEAST YHR131C.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z46843; CAAG6882.1; -  
DR EMBL: X92517; CAAG63294.1; -  
DR EMBL: Z71420; CAAG6027.1; -  
DR SGD: S0005088; YNL144C.  
KW Hypothetical protein.  
SQ SEQUENCE 740 AA; 84202 MW; 7A3D75999DC83D06 CRC64;

Query Match 52.0%; Score 39; DB 1; Length 740;  
Best Local Similarity 53.3%; Pred. No. 53;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSIICSP--PSLSLS 15  
Db 42 SSSIICSP--PSLSLS 56

RESULT 9  
YNE6\_CAEEL STANDARD: PRT: 918 AA.  
ID YNE6\_CAEEL STANDARD: PRT: 918 AA.  
AC P30645;  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable 3',5'-cyclic phosphodiesterase R08D7.6 (EC 3.1.4.17).  
GN R08D7.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Lareelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smailton N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
CC -1- CATALYTIC ACTIVITY: Nucleoside 3',5'-cyclic phosphate + H(2)O =  
CC nucleoside 5'-phosphate.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
CC FAMILY. STRONG, TO MAMMALIAN TYPE 6 GMP PHOSPHODIESTERASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z12017; CAAT8052.1; -  
DR PIR: S24462; S24462.  
DR PIR: S41041; S41041.  
DR WormPep: R08D7.6; CE01047.  
DR InterPro: IPR003018; GAF.  
DR InterPro: IPR003607; ME\_PpIase\_HDC.  
DR InterPro: IPR002073; PDEase.  
DR Pfam: PF00233; PDEase; 1.  
DR Pfam: PF01590; GAF; 1.  
DR PRINTS: PRO0387; PDIESTERASE1.  
DR SMART: SM00065; GAF; 1.  
DR SMART: SM00471; HDC; 1.  
DR PROSITE: PS00126; PDEASE\_I; 1.  
KW Hypothetical protein; Hydrolyase; GMP.  
SQ SEQUENCE 918 AA; 104238 MW; CA59910AF4CD6155 CRC64;

Query Match 52.0%; Score 39; DB 1; Length 918;  
Best Local Similarity 72.7%; Pred. No. 67;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;



OX NCBI\_TaxID=9940;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-Pituitary;  
 RX MEDLINE=94059099; PubMed=7694577;  
 RA ILLING N., Jacobs G.F.M., Becker I.I., Flanagan C.A., Davidson J.S.,  
 RA Eales A., Zhou W., Sealton S.C., Millar R.P.;  
 RT "Comparative sequence analysis and functional characterization of the  
 RT cloned sheep gonadotropin-releasing hormone receptor reveal  
 RT differences in primary structure and ligand specificity among  
 RT mammalian receptors";  
 RL Biochem. Biophys. Res. Commun. 196:745-751(1993).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-Scottish blackface; TISSUE-Pituitary;  
 RX MEDLINE=94040345; PubMed=8224516;  
 RA Brooks J., Taylor P.L., Saunders P.T.K., Eide K.A., Scruthers W.J.,  
 RA McNeilly A.S.;  
 RT "Cloning and sequencing of the sheep pituitary gonadotropin-releasing  
 RT hormone receptor and changes in expression of its mRNA during the  
 RT estrous cycle.";  
 RL Mol. Cell. Endocrinol. 94:R23-R27(1993).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE-Liver;  
 RX MEDLINE=96235149; PubMed=8666259;  
 RA Campion C.E., Turzillo A.M., Clay C.M.;  
 RT "The gene encoding the ovine gonadotropin-releasing hormone (GnRH)  
 RT receptor: cloning and initial characterization.";  
 RL Gene 170:277-280(1996).  
 CC -1- FUNCTION: THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH  
 CC G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND  
 CC MESSENGER SYSTEM.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb-sib.ch](mailto:license@sdb-sib.ch)).  
 CC -----  
 DR EMBL: L22215; AAC37336.1; -  
 DR EMBL: X72088; CA50978.1; -  
 DR EMBL: L42937; AAB38515.1; -  
 DR EMBL: L43842; ABA41939.1; -  
 DR EMBL: L43841; ABA41939.1; JOINED.  
 DR PIR: JN0882; JN0882.  
 DR InterPro: IPR000276; GPCR\_Rhodpsn.  
 DR Pfam: PF00001; 7tm\_1; 1.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPT\_FL\_1; 1.  
 DR PROSITE: PS50262; G\_PROTEIN\_RECPT\_FL\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1  
 FT TRANSMEM 39  
 FT TRANSMEM 58  
 FT TRANSMEM 77  
 FT TRANSMEM 97  
 FT TRANSMEM 98  
 FT TRANSMEM 115  
 FT TRANSMEM 116  
 FT TRANSMEM 137  
 FT TRANSMEM 138  
 FT TRANSMEM 164  
 FT TRANSMEM 165  
 FT TRANSMEM 184  
 FT TRANSMEM 185  
 FT TRANSMEM 212  
 FT TRANSMEM 213  
 FT TRANSMEM 232  
 FT TRANSMEM 233  
 FT TRANSMEM 281  
 FT TRANSMEM 282  
 FT TRANSMEM 300  
 FT TRANSMEM 301  
 FT TRANSMEM 306  
 FT TRANSMEM 307  
 FT TRANSMEM 326  
 FT TRANSMEM 327  
 FT TRANSMEM 328  
 FT TRANSMEM 372  
 FT TRANSMEM 373  
 FT TRANSMEM 374  
 FT TRANSMEM 375  
 FT TRANSMEM 376  
 FT TRANSMEM 377  
 FT TRANSMEM 378  
 FT TRANSMEM 379  
 FT TRANSMEM 380  
 FT TRANSMEM 381  
 FT TRANSMEM 382  
 FT TRANSMEM 383  
 FT TRANSMEM 384  
 FT TRANSMEM 385  
 FT TRANSMEM 386  
 FT TRANSMEM 387  
 FT TRANSMEM 388  
 FT TRANSMEM 389  
 FT TRANSMEM 390  
 FT TRANSMEM 391  
 FT TRANSMEM 392  
 FT TRANSMEM 393  
 FT TRANSMEM 394  
 FT TRANSMEM 395  
 FT TRANSMEM 396  
 FT TRANSMEM 397  
 FT TRANSMEM 398  
 FT TRANSMEM 399  
 FT TRANSMEM 400  
 FT TRANSMEM 401  
 FT TRANSMEM 402  
 FT TRANSMEM 403  
 FT TRANSMEM 404  
 FT TRANSMEM 405  
 FT TRANSMEM 406  
 FT TRANSMEM 407  
 FT TRANSMEM 408  
 FT TRANSMEM 409  
 FT TRANSMEM 410  
 FT TRANSMEM 411  
 FT TRANSMEM 412  
 FT TRANSMEM 413  
 FT TRANSMEM 414  
 FT TRANSMEM 415  
 FT TRANSMEM 416  
 FT TRANSMEM 417  
 FT TRANSMEM 418  
 FT TRANSMEM 419  
 FT TRANSMEM 420  
 FT TRANSMEM 421  
 FT TRANSMEM 422  
 FT TRANSMEM 423  
 FT TRANSMEM 424  
 FT TRANSMEM 425  
 FT TRANSMEM 426  
 FT TRANSMEM 427  
 FT TRANSMEM 428  
 FT TRANSMEM 429  
 FT TRANSMEM 430  
 FT TRANSMEM 431  
 FT TRANSMEM 432  
 FT TRANSMEM 433  
 FT TRANSMEM 434  
 FT TRANSMEM 435  
 FT TRANSMEM 436  
 FT TRANSMEM 437  
 FT TRANSMEM 438  
 FT TRANSMEM 439  
 FT TRANSMEM 440  
 FT TRANSMEM 441  
 FT TRANSMEM 442  
 FT TRANSMEM 443  
 FT TRANSMEM 444  
 FT TRANSMEM 445  
 FT TRANSMEM 446  
 FT TRANSMEM 447  
 FT TRANSMEM 448  
 FT TRANSMEM 449  
 FT TRANSMEM 450  
 FT TRANSMEM 451  
 FT TRANSMEM 452  
 FT TRANSMEM 453  
 FT TRANSMEM 454  
 FT TRANSMEM 455  
 FT TRANSMEM 456  
 FT TRANSMEM 457  
 FT TRANSMEM 458  
 FT TRANSMEM 459  
 FT TRANSMEM 460  
 FT TRANSMEM 461  
 FT TRANSMEM 462  
 FT TRANSMEM 463  
 FT TRANSMEM 464  
 FT TRANSMEM 465  
 FT TRANSMEM 466  
 FT TRANSMEM 467  
 FT TRANSMEM 468  
 FT TRANSMEM 469  
 FT TRANSMEM 470  
 FT TRANSMEM 471  
 FT TRANSMEM 472  
 FT TRANSMEM 473  
 FT TRANSMEM 474  
 FT TRANSMEM 475  
 FT TRANSMEM 476  
 FT TRANSMEM 477  
 FT TRANSMEM 478  
 FT TRANSMEM 479  
 FT TRANSMEM 480  
 FT TRANSMEM 481  
 FT TRANSMEM 482  
 FT TRANSMEM 483  
 FT TRANSMEM 484  
 FT TRANSMEM 485  
 FT TRANSMEM 486  
 FT TRANSMEM 487  
 FT TRANSMEM 488  
 FT TRANSMEM 489  
 FT TRANSMEM 490  
 FT TRANSMEM 491  
 FT TRANSMEM 492  
 FT TRANSMEM 493  
 FT TRANSMEM 494  
 FT TRANSMEM 495  
 FT TRANSMEM 496  
 FT TRANSMEM 497  
 FT TRANSMEM 498  
 FT TRANSMEM 499  
 FT TRANSMEM 500  
 FT TRANSMEM 501  
 FT TRANSMEM 502  
 FT TRANSMEM 503  
 FT TRANSMEM 504  
 FT TRANSMEM 505  
 FT TRANSMEM 506  
 FT TRANSMEM 507  
 FT TRANSMEM 508  
 FT TRANSMEM 509  
 FT TRANSMEM 510  
 FT TRANSMEM 511  
 FT TRANSMEM 512  
 FT TRANSMEM 513  
 FT TRANSMEM 514  
 FT TRANSMEM 515  
 FT TRANSMEM 516  
 FT TRANSMEM 517  
 FT TRANSMEM 518  
 FT TRANSMEM 519  
 FT TRANSMEM 520  
 FT TRANSMEM 521  
 FT TRANSMEM 522  
 FT TRANSMEM 523  
 FT TRANSMEM 524  
 FT TRANSMEM 525  
 FT TRANSMEM 526  
 FT TRANSMEM 527  
 FT TRANSMEM 528  
 FT TRANSMEM 529  
 FT TRANSMEM 530  
 FT TRANSMEM 531  
 FT TRANSMEM 532  
 FT TRANSMEM 533  
 FT TRANSMEM 534  
 FT TRANSMEM 535  
 FT TRANSMEM 536  
 FT TRANSMEM 537  
 FT TRANSMEM 538  
 FT TRANSMEM 539  
 FT TRANSMEM 540  
 FT TRANSMEM 541  
 FT TRANSMEM 542  
 FT TRANSMEM 543  
 FT TRANSMEM 544  
 FT TRANSMEM 545  
 FT TRANSMEM 546  
 FT TRANSMEM 547  
 FT TRANSMEM 548  
 FT TRANSMEM 549  
 FT TRANSMEM 550  
 FT TRANSMEM 551  
 FT TRANSMEM 552  
 FT TRANSMEM 553  
 FT TRANSMEM 554  
 FT TRANSMEM 555  
 FT TRANSMEM 556  
 FT TRANSMEM 557  
 FT TRANSMEM 558  
 FT TRANSMEM 559  
 FT TRANSMEM 560  
 FT TRANSMEM 561  
 FT TRANSMEM 562  
 FT TRANSMEM 563  
 FT TRANSMEM 564  
 FT TRANSMEM 565  
 FT TRANSMEM 566  
 FT TRANSMEM 567  
 FT TRANSMEM 568  
 FT TRANSMEM 569  
 FT TRANSMEM 570  
 FT TRANSMEM 571  
 FT TRANSMEM 572  
 FT TRANSMEM 573  
 FT TRANSMEM 574  
 FT TRANSMEM 575  
 FT TRANSMEM 576  
 FT TRANSMEM 577  
 FT TRANSMEM 578  
 FT TRANSMEM 579  
 FT TRANSMEM 580  
 FT TRANSMEM 581  
 FT TRANSMEM 582  
 FT TRANSMEM 583  
 FT TRANSMEM 584  
 FT TRANSMEM 585  
 FT TRANSMEM 586  
 FT TRANSMEM 587  
 FT TRANSMEM 588  
 FT TRANSMEM 589  
 FT TRANSMEM 590  
 FT TRANSMEM 591  
 FT TRANSMEM 592  
 FT TRANSMEM 593  
 FT TRANSMEM 594  
 FT TRANSMEM 595  
 FT TRANSMEM 596  
 FT TRANSMEM 597  
 FT TRANSMEM 598  
 FT TRANSMEM 599  
 FT TRANSMEM 600  
 FT TRANSMEM 601  
 FT TRANSMEM 602  
 FT TRANSMEM 603  
 FT TRANSMEM 604  
 FT TRANSMEM 605  
 FT TRANSMEM 606  
 FT TRANSMEM 607  
 FT TRANSMEM 608  
 FT TRANSMEM 609  
 FT TRANSMEM 610  
 FT TRANSMEM 611  
 FT TRANSMEM 612  
 FT TRANSMEM 613  
 FT TRANSMEM 614  
 FT TRANSMEM 615  
 FT TRANSMEM 616  
 FT TRANSMEM 617  
 FT TRANSMEM 618  
 FT TRANSMEM 619  
 FT TRANSMEM 620  
 FT TRANSMEM 621  
 FT TRANSMEM 622  
 FT TRANSMEM 623  
 FT TRANSMEM 624  
 FT TRANSMEM 625  
 FT TRANSMEM 626  
 FT TRANSMEM 627  
 FT TRANSMEM 628  
 FT TRANSMEM 629  
 FT TRANSMEM 630  
 FT TRANSMEM 631  
 FT TRANSMEM 632  
 FT TRANSMEM 633  
 FT TRANSMEM 634  
 FT TRANSMEM 635  
 FT TRANSMEM 636  
 FT TRANSMEM 637  
 FT TRANSMEM 638  
 FT TRANSMEM 639  
 FT TRANSMEM 640  
 FT TRANSMEM 641  
 FT TRANSMEM 642  
 FT TRANSMEM 643  
 FT TRANSMEM 644  
 FT TRANSMEM 645  
 FT TRANSMEM 646  
 FT TRANSMEM 647  
 FT TRANSMEM 648  
 FT TRANSMEM 649  
 FT TRANSMEM 650  
 FT TRANSMEM 651  
 FT TRANSMEM 652  
 FT TRANSMEM 653  
 FT TRANSMEM 654  
 FT TRANSMEM 655  
 FT TRANSMEM 656  
 FT TRANSMEM 657  
 FT TRANSMEM 658  
 FT TRANSMEM 659  
 FT TRANSMEM 660  
 FT TRANSMEM 661  
 FT TRANSMEM 662  
 FT TRANSMEM 663  
 FT TRANSMEM 664  
 FT TRANSMEM 665  
 FT TRANSMEM 666  
 FT TRANSMEM 667  
 FT TRANSMEM 668  
 FT TRANSMEM 669  
 FT TRANSMEM 670  
 FT TRANSMEM 671  
 FT TRANSMEM 672  
 FT TRANSMEM 673  
 FT TRANSMEM 674  
 FT TRANSMEM 675  
 FT TRANSMEM 676  
 FT TRANSMEM 677  
 FT TRANSMEM 678  
 FT TRANSMEM 679  
 FT TRANSMEM 680  
 FT TRANSMEM 681  
 FT TRANSMEM 682  
 FT TRANSMEM 683  
 FT TRANSMEM 684  
 FT TRANSMEM 685  
 FT TRANSMEM 686  
 FT TRANSMEM 687  
 FT TRANSMEM 688  
 FT TRANSMEM 689  
 FT TRANSMEM 690  
 FT TRANSMEM 691  
 FT TRANSMEM 692  
 FT TRANSMEM 693  
 FT TRANSMEM 694  
 FT TRANSMEM 695  
 FT TRANSMEM 696  
 FT TRANSMEM 697  
 FT TRANSMEM 698  
 FT TRANSMEM 699  
 FT TRANSMEM 700  
 FT TRANSMEM 701  
 FT TRANSMEM 702  
 FT TRANSMEM 703  
 FT TRANSMEM 704  
 FT TRANSMEM 705  
 FT TRANSMEM 706  
 FT TRANSMEM 707  
 FT TRANSMEM 708  
 FT TRANSMEM 709  
 FT TRANSMEM 710  
 FT TRANSMEM 711  
 FT TRANSMEM 712  
 FT TRANSMEM 713  
 FT TRANSMEM 714  
 FT TRANSMEM 715  
 FT TRANSMEM 716  
 FT TRANSMEM 717  
 FT TRANSMEM 718  
 FT TRANSMEM 719  
 FT TRANSMEM 720  
 FT TRANSMEM 721  
 FT TRANSMEM 722  
 FT TRANSMEM 723  
 FT TRANSMEM 724  
 FT TRANSMEM 725  
 FT TRANSMEM 726  
 FT TRANSMEM 727  
 FT TRANSMEM 728  
 FT TRANSMEM 729  
 FT TRANSMEM 730  
 FT TRANSMEM 731  
 FT TRANSMEM 732  
 FT TRANSMEM 733  
 FT TRANSMEM 734  
 FT TRANSMEM 735  
 FT TRANSMEM 736  
 FT TRANSMEM 737  
 FT TRANSMEM 738  
 FT TRANSMEM 739  
 FT TRANSMEM 740  
 FT TRANSMEM 741  
 FT TRANSMEM 742  
 FT TRANSMEM 743  
 FT TRANSMEM 744  
 FT TRANSMEM 745  
 FT TRANSMEM 746  
 FT TRANSMEM 747  
 FT TRANSMEM 748  
 FT TRANSMEM 749  
 FT TRANSMEM 750  
 FT TRANSMEM 751  
 FT TRANSMEM 752  
 FT TRANSMEM 753  
 FT TRANSMEM 754  
 FT TRANSMEM 755  
 FT TRANSMEM 756  
 FT TRANSMEM 757  
 FT TRANSMEM 758  
 FT TRANSMEM 759  
 FT TRANSMEM 760  
 FT TRANSMEM 761  
 FT TRANSMEM 762  
 FT TRANSMEM 763  
 FT TRANSMEM 764  
 FT TRANSMEM 765  
 FT TRANSMEM 766  
 FT TRANSMEM 767  
 FT TRANSMEM 768  
 FT TRANSMEM 769  
 FT TRANSMEM 770  
 FT TRANSMEM 771  
 FT TRANSMEM 772  
 FT TRANSMEM 773  
 FT TRANSMEM 774  
 FT TRANSMEM 775  
 FT TRANSMEM 776  
 FT TRANSMEM 777  
 FT TRANSMEM 778  
 FT TRANSMEM 779  
 FT TRANSMEM 780  
 FT TRANSMEM 781  
 FT TRANSMEM 782  
 FT TRANSMEM 783  
 FT TRANSMEM 784  
 FT TRANSMEM 785  
 FT TRANSMEM 786  
 FT TRANSMEM 787  
 FT TRANSMEM 788  
 FT TRANSMEM 789  
 FT TRANSMEM 790  
 FT TRANSMEM 791  
 FT TRANSMEM 792  
 FT TRANSMEM 793  
 FT TRANSMEM 794  
 FT TRANSMEM 795  
 FT TRANSMEM 796  
 FT TRANSMEM 797  
 FT TRANSMEM 798  
 FT TRANSMEM 799  
 FT TRANSMEM 800  
 FT TRANSMEM 801  
 FT TRANSMEM 802  
 FT TRANSMEM 803  
 FT TRANSMEM 804  
 FT TRANSMEM 805  
 FT TRANSMEM 806  
 FT TRANSMEM 807  
 FT TRANSMEM 808  
 FT TRANSMEM 809  
 FT TRANSMEM 810  
 FT TRANSMEM 811  
 FT TRANSMEM 812  
 FT TRANSMEM 813  
 FT TRANSMEM 814  
 FT TRANSMEM 815  
 FT TRANSMEM 816  
 FT TRANSMEM 817  
 FT TRANSMEM 818  
 FT TRANSMEM 819  
 FT TRANSMEM 820  
 FT TRANSMEM 821  
 FT TRANSMEM 822  
 FT TRANSMEM 823  
 FT TRANSMEM 824  
 FT TRANSMEM 825  
 FT TRANSMEM 826  
 FT TRANSMEM 827  
 FT TRANSMEM 828  
 FT TRANSMEM 829  
 FT TRANSMEM 830  
 FT TRANSMEM 831  
 FT TRANSMEM 832  
 FT TRANSMEM 833  
 FT TRANSMEM 834  
 FT TRANSMEM 835  
 FT TRANSMEM 836  
 FT TRANSMEM 837  
 FT TRANSMEM 838  
 FT TRANSMEM 839  
 FT TRANSMEM 840  
 FT TRANSMEM 841  
 FT TRANSMEM 842  
 FT TRANSMEM 843  
 FT TRANSMEM 844  
 FT TRANSMEM 845  
 FT TRANSMEM 846  
 FT TRANSMEM 847  
 FT TRANSMEM 848  
 FT TRANSMEM 849  
 FT TRANSMEM 850  
 FT TRANSMEM 851  
 FT TRANSMEM 852  
 FT TRANSMEM 853  
 FT TRANSMEM 854  
 FT TRANSMEM 855  
 FT TRANSMEM 856  
 FT TRANSMEM 857  
 FT TRANSMEM 858  
 FT TRANSMEM 859  
 FT TRANSMEM 860  
 FT TRANSMEM 861  
 FT TRANSMEM 862  
 FT TRANSMEM 863  
 FT TRANSMEM 864  
 FT TRANSMEM 865  
 FT TRANSMEM 866  
 FT TRANSMEM 867  
 FT TRANSMEM 868  
 FT TRANSMEM 869  
 FT TRANSMEM 870  
 FT TRANSMEM 871  
 FT TRANSMEM 872  
 FT TRANSMEM 873  
 FT TRANSMEM 874  
 FT TRANSMEM 875  
 FT TRANSMEM 876  
 FT TRANSMEM 877  
 FT TRANSMEM 878  
 FT TRANSMEM 879  
 FT TRANSMEM 880  
 FT TRANSMEM 881  
 FT TRANSMEM 882  
 FT TRANSMEM 883  
 FT TRANSMEM 884  
 FT TRANSMEM 885  
 FT TRANSMEM 886  
 FT TRANSMEM 887  
 FT TRANSMEM 888  
 FT TRANSMEM 889  
 FT TRANSMEM 890  
 FT TRANSMEM 891  
 FT TRANSMEM 892  
 FT TRANSMEM 893  
 FT TRANSMEM 894  
 FT TRANSMEM 895  
 FT TRANSMEM 896  
 FT TRANSMEM 897  
 FT TRANSMEM 898  
 FT TRANSMEM 899  
 FT TRANSMEM 900  
 FT TRANSMEM 901  
 FT TRANSMEM 902  
 FT TRANSMEM 903  
 FT TRANSMEM 904  
 FT TRANSMEM 905  
 FT TRANSMEM 906  
 FT TRANSMEM 907  
 FT TRANSMEM 908  
 FT TRANSMEM 909  
 FT TRANSMEM 910  
 FT TRANSMEM 911  
 FT TRANSMEM 912  
 FT TRANSMEM 913  
 FT TRANSMEM 914  
 FT TRANSMEM 915  
 FT TRANSMEM 916  
 FT TRANSMEM 917  
 FT TRANSMEM 918  
 FT TRANSMEM 919  
 FT TRANSMEM 920  
 FT TRANSMEM 921  
 FT TRANSMEM 922  
 FT TRANSMEM 923  
 FT TRANSMEM 924  
 FT TRANSMEM 925  
 FT TRANSMEM 926  
 FT TRANSMEM 927  
 FT TRANSMEM 928  
 FT TRANSMEM 929  
 FT TRANSMEM 930  
 FT TRANSMEM 931  
 FT TRANSMEM 932  
 FT TRANSMEM 933  
 FT TRANSMEM 934  
 FT TRANSMEM 935  
 FT TRANSMEM 936  
 FT TRANSMEM 937  
 FT TRANSMEM 938  
 FT TRANSMEM 939  
 FT TRANSMEM 940  
 FT TRANSMEM 941  
 FT TRANSMEM 942  
 FT TRANSMEM 943  
 FT TRANSMEM 944  
 FT TRANSMEM 945  
 FT TRANSMEM 946  
 FT TRANSMEM 947  
 FT TRANSMEM 948  
 FT TRANSMEM 949  
 FT TRANSMEM 950  
 FT TRANSMEM 951  
 FT TRANSMEM 952  
 FT TRANSMEM 953  
 FT TRANSMEM 954  
 FT TRANSMEM 955  
 FT TRANSMEM 956  
 FT TRANSMEM 957  
 FT TRANSMEM 958  
 FT TRANSMEM 959  
 FT TRANSMEM 960  
 FT TRANSMEM 961  
 FT TRANSMEM 962  
 FT TRANSMEM 963  
 FT TRANSMEM 964  
 FT TRANSMEM 965  
 FT TRANSMEM 966  
 FT TRANSMEM 967  
 FT TRANSMEM 968  
 FT TRANSMEM 969  
 FT TRANSMEM 970  
 FT TRANSMEM 971  
 FT TRANSMEM 972  
 FT TRANSMEM 973  
 FT TRANSMEM 974  
 FT TRANSMEM 975  
 FT TRANSMEM 976  
 FT TRANSMEM 977  
 FT TRANSMEM 978  
 FT TRANSMEM 979  
 FT TRANSMEM 980  
 FT TRANSMEM 981  
 FT TRANSMEM 982  
 FT TRANSMEM 983  
 FT TRANSMEM 984  
 FT TRANSMEM 985  
 FT TRANSMEM 986  
 FT TRANSMEM 987  
 FT TRANSMEM 988  
 FT TRANSMEM 989  
 FT TRANSMEM 990  
 FT TRANSMEM 991  
 FT TRANSMEM 992  
 FT TRANSMEM 993  
 FT TRANSMEM 994  
 FT TRANSMEM 995  
 FT TRANSMEM 996  
 FT TRANSMEM 997  
 FT TRANSMEM 998  
 FT TRANSMEM 999  
 FT TRANSMEM 1000  
 FT TRANSMEM 1001  
 FT TRANSMEM 1002  
 FT TRANSMEM 1003  
 FT TRANSMEM 1004  
 FT TRANSMEM 1005  
 FT TRANSMEM 1006  
 FT TRANSMEM 1007  
 FT TRANSMEM 1008  
 FT TRANSMEM 1009  
 FT TRANSMEM 1010  
 FT TRANSMEM 1011  
 FT TRANSMEM 1012  
 FT TRANSMEM 1013  
 FT TRANSMEM 1014  
 FT TRANSMEM 1015  
 FT TRANSMEM 1016  
 FT TRANSMEM 1017  
 FT TRANSMEM 1018  
 FT TRANSMEM 1019  
 FT TRANSMEM 1020  
 FT TRANSMEM 1021  
 FT TRANSMEM 1022  
 FT TRANSMEM 1023  
 FT TRANSMEM 1024  
 FT TRANSMEM 1025  
 FT TRANSMEM 1026  
 FT TRANSMEM 1027  
 FT TRANSMEM 1028  
 FT TRANSMEM 1029  
 FT TRANSMEM 1030  
 FT TRANSMEM 1031  
 FT TRANSMEM 1032  
 FT TRANSMEM 1033  
 FT TRANSMEM 1034  
 FT TRANSMEM 1035  
 FT TRANSMEM 1036  
 FT TRANSMEM 1037  
 FT TRANSMEM 1038  
 FT TRANSMEM 1039  
 FT TRANSMEM 1040  
 FT TRANSMEM 1041  
 FT TRANSMEM 1042  
 FT TRANSMEM 1043  
 FT TRANSMEM 1044  
 FT TRANSMEM 1045  
 FT TRANSMEM 1046  
 FT TRANSMEM 1047  
 FT TRANSMEM 1048  
 FT TRANSMEM 1049  
 FT TRANSMEM 1050  
 FT TRANSMEM 1051  
 FT TRANSMEM 1052  
 FT TRANSMEM 1053  
 FT TRANSMEM 1054  
 FT TRANSMEM 1055  
 FT TRANSMEM 1056  
 FT TRANSMEM 1057  
 FT TRANSMEM 1058  
 FT TRANSMEM 1059  
 FT TRANSMEM 1060  
 FT TRANSMEM 1061  
 FT TRANSMEM 1062  
 FT TRANSMEM 1063  
 FT TRANSMEM 1064  
 FT TRANSMEM 1065  
 FT TRANSMEM 1066  
 FT TRANSMEM 1067  
 FT TRANSMEM 1068  
 FT TRANSMEM 1069  
 FT TRANSMEM 1070  
 FT TRANSMEM 1071  
 FT TRANSMEM 1072  
 FT TRANSMEM 1073  
 FT TRANSMEM 1074  
 FT TRANSMEM 1075  
 FT TRANSMEM 1076  
 FT TRANSMEM 1077  
 FT TRANSMEM 1078  
 FT TRANSMEM 1079  
 FT TRANSMEM 1080  
 FT TRANSMEM 1081  
 FT TRANSMEM 1082  
 FT TRANSMEM 1083  
 FT TRANSMEM 1084  
 FT TRANSMEM 1085  
 FT TRANSMEM 1086  
 FT TRANSMEM 1087  
 FT TRANSMEM 1088  
 FT TRANSMEM 1089  
 FT TRANSMEM 1090  
 FT TRANSMEM 1091  
 FT TRANSMEM 1092  
 FT TRANSMEM 1093  
 FT TRANSMEM 1094  
 FT TRANSMEM 1095  
 FT TRANSMEM 1096  
 FT TRANSMEM 1097  
 FT TRANSMEM 1098  
 FT TRANSMEM 1099  
 FT TRANSMEM 1100  
 FT TRANSMEM 1101  
 FT TRANSMEM 1102  
 FT TRANSMEM 1103  
 FT TRANSMEM 1104  
 FT TRANSMEM 1105  
 FT TRANSMEM 1106  
 FT TRANSMEM 1107  
 FT TRANSMEM 1108  
 FT TRANSMEM 1109  
 FT TRANSMEM 1110  
 FT TRANSMEM 1111  
 FT TRANSMEM 1112  
 FT TRANSMEM 1113  
 FT TRANSMEM 1114  
 FT TRANSMEM 1115  
 FT TRANSMEM 1116  
 FT TRANSMEM 1117  
 FT TRANSMEM 1118  
 FT TRANSMEM 1119  
 FT TRANSMEM 1120  
 FT TRANSMEM 1121  
 FT TRANSMEM 1122  
 FT TRANSMEM 1123  
 FT TRANSMEM 1124  
 FT TRANSMEM 1125  
 FT TRANSMEM 1126  
 FT TRANSMEM 1127  
 FT TRANSMEM 1128  
 FT TRANSMEM 1129  
 FT TRANSMEM 1130  
 FT TRANSMEM 1131  
 FT TRANSMEM 1132  
 FT TRANSMEM 1133  
 FT TRANSMEM 1134  
 FT TRANSMEM 1135  
 FT TRANSMEM 1136  
 FT TRANSMEM 1137  
 FT TRANSMEM 1138  
 FT TRANSMEM 1139  
 FT TRANSMEM 1140  
 FT TRANSMEM 1141  
 FT TRANSMEM 1142  
 FT TRANSMEM 1143  
 FT TRANSMEM 1144  
 FT TRANSMEM 1145  
 FT TRANSMEM 1146  
 FT TRANSMEM 1147  
 FT TRANSMEM 1148  
 FT TRANSMEM 1149  
 FT TRANSMEM 1150  
 FT TRANSMEM 1151  
 FT TRANSMEM 1152  
 FT TRANSMEM 1153  
 FT TRANSMEM 1154  
 FT TRANSMEM 1155  
 FT TRANSMEM 1156  
 FT TRANSMEM 1157  
 FT TRANSMEM 1158  
 FT TRANSMEM 1159  
 FT TRANSMEM 1160  
 FT TRANSMEM 1161  
 FT TRANSMEM 1162  
 FT TRANSMEM 1163  
 FT TRANSMEM 1164  
 FT TRANSMEM 1165  
 FT TRANSMEM 1166  
 FT TRANSMEM 1167  
 FT TRANSMEM 1168  
 FT TRANSMEM 1169  
 FT TRANSMEM 1170  
 FT TRANSMEM 1171  
 FT TRANSMEM 1172  
 FT TRANSMEM 1173  
 FT TRANSMEM 1174  
 FT TRANSMEM 1175  
 FT TRANSMEM 1176  
 FT TRANSMEM 1177  
 FT TRANSMEM 1178  
 FT TRANSMEM 1179  
 FT TRANSMEM 1180  
 FT TRANSMEM 1181  
 FT TRANSMEM 1182  
 FT TRANSMEM 1183  
 FT TRANSMEM 1184  
 FT TRANSMEM 1185  
 FT TRANSMEM 1186  
 FT TRANSMEM 1187  
 FT TRANSMEM 1188  
 FT TRANSMEM 1189  
 FT TRANSMEM 1190  
 FT TRANSMEM 1191  
 FT TRANSMEM 1192  
 FT TRANSMEM 1193  
 FT TRANSMEM 1194  
 FT TRANSMEM 1195  
 FT TRANSMEM 1196  
 FT TRANSMEM 1197  
 FT TRANSMEM 1198  
 FT TRANSMEM 1199  
 FT TRANSMEM 1200  
 FT TRANSMEM 1201  
 FT TRANSMEM 1202  
 FT TRANSMEM 1203  
 FT TRANSMEM 1204  
 FT TRANSMEM 1205  
 FT TRANSMEM 1206  
 FT TRANSMEM 1207  
 FT TRANSMEM 1208  
 FT TRANSMEM 1209  
 FT TRANSMEM 1210  
 FT TRANSMEM 1211  
 FT TRANSMEM 1212  
 FT TRANSMEM 1213  
 FT TRANSMEM 1214  
 FT TRANSMEM 1215  
 FT TRANSMEM 1216  
 FT TRANSMEM 1217  
 FT TRANSMEM 1218  
 FT TRANSMEM 1219  
 FT TRANSMEM 1220  
 FT TRANSMEM 1221  
 FT TRANSMEM 1222  
 FT TRANSMEM 1223  
 FT TRANSMEM 1224  
 FT TRANSMEM 1225  
 FT TRANSMEM 1226  
 FT TRANSMEM 1227  
 FT TRANSMEM 1228  
 FT TRANSMEM 1229  
 FT TRANSMEM 1230  
 FT TRANSMEM 1231  
 FT TRANSMEM 1232  
 FT TRANSMEM 1233  
 FT TRANSMEM 1234  
 FT TRANSMEM 1235  
 FT TRANSMEM 1236  
 FT TRANSMEM 1237  
 FT TRANSMEM 1238  
 FT TRANSMEM 1239  
 FT TRANSMEM 1240  
 FT TRANSMEM 1241  
 FT TRANSMEM 1242  
 FT TRANSMEM 1243  
 FT TRANSMEM 1244  
 FT TRANSMEM 1245  
 FT TRANSMEM 1246  
 FT TRANSMEM 1247  
 FT TRANSMEM 1248  
 FT TRANSM





```
CC      pyruvate + L-glutamate.
CC      -1- PATHWAY: Tryptophan biosynthesis; first step.
CC      -1- SUBUNIT: TETRAMER OF TWO COMPONENTS I AND TWO COMPONENTS II.
CC      -1- SUBCELLULAR LOCATION: Chloroplast (Probable).
CC      -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE
CC      USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES
CC      GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
CC      -1- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
CC      FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M92354; AAA32739.1; -.
DR      EMBL; AC005496; AAC35228.1; -.
DR      PIR; S27752; S27752.
DR      HSSP; Q06128; 10DL.
DR      InterPro: IPR005256; Anth_synthI.
DR      InterPro: IPR000350; Chorismate_bind.
DR      Pfam; PF00425; Chorismate_bind; 1.
DR      PRINTS; PR00095; ANTSNTHASEI.
DR      ProDom; PD000779; Chorismate_bind; 1.
DR      TIGRFAMs; TIGR00564; trpE_most; 1.
KW      Tryptophan biosynthesis; Lyase; Chloroplast; Transit peptide;
KW      Multigene family.
FT      TRANSIT 1 ?
FT      CHAIN 1 621 CHLOROPLAST (POTENTIAL).
SQ      SEQUENCE 621 AA; 69815 MW; DA039FE58A420967 CRC64;
      ANTHRANILATE SYNTHASE COMPONENT I-2.

Query Match      50.7%; Score 38; DB 1; Length 621;
Best Local Similarity 59.2%; Pred. No. 65;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy      3 SILCSPLPSLSLS 15
      |||||
db      66 SIKCSYTPSLDLS 78
```

Search completed: June 25, 2003, 17:15:58  
Job time : 2.6443 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: June 25, 2003, 17:13:30 ; Search time 1.93289 Seconds  
(without alignments)  
1705.612 Million cell updates/sec

Title: US-10-075-846-4\_COPY\_192\_207

Perfect score: 75  
Sequence: 1 SSSILCSPPLSLSLV 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp-archaea:\*
- 2: sp-bacteria:\*
- 3: sp-fungi:\*
- 4: sp-human:\*
- 5: sp-invertebrate:\*
- 6: sp-mammal:\*
- 7: sp-mhc:\*
- 8: sp-organelle:\*
- 9: sp-phage:\*
- 10: sp-plant:\*
- 11: sp-rodent:\*
- 12: sp-virus:\*
- 13: sp-vertebrate:\*
- 14: sp-unclassified:\*
- 15: sp-virus:\*
- 16: sp\_bacteriap:\*
- 17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	65.3	490	10 Q9SUV8	Q9SUV8 arabidopsis
2	47	62.7	1298	11 Q8R4H6	Q8R4H6 mus musculu
3	46	61.3	94	16 Q9JMS9	Q9JMS9 escherichia
4	46	61.3	117	16 Q8X3Q9	Q8X3Q9 escherichia
5	46	61.3	490	10 Q9MAA0	Q9MAA0 arabidopsis
6	46	61.3	555	10 Q8W0Z7	Q8W0Z7 arabidopsis
7	46	61.3	724	11 Q91ZA3	Q91ZA3 mus musculu
8	46	61.3	724	11 Q922N3	Q922N3 mus musculu
9	44	58.7	887	4 Q15134	Q15134 homo sapien
10	43	57.3	82	9 Q8W6E1	Q8W6E1 bacterioph
11	43	57.3	138	5 Q9N7D6	Q9N7D6 leishmania
12	43	57.3	189	4 Q96EH4	Q96EH4 homo sapien
13	43	57.3	3567	11 Q9ES77	Q9ES77 mus musculu
14	42	56.0	113	11 Q9QWV3	Q9QWV3 mus musculu
15	42	56.0	114	17 Q9YBN2	Q9YBN2 aeropyrum p
16	42	56.0	123	4 Q8WYV1	Q8WYV1 homo sapien

17	42	56.0	141	16 Q9PH15	Q9PH15 xylella fas
18	42	56.0	478	16 Q83294	Q83294 treponema p
19	42	56.0	739	11 Q35613	Q35613 mus musculu
20	42	56.0	740	11 Q9QWV8	Q9QWV8 mus musculu
21	42	56.0	1010	3 Q9USN1	Q9USN1 schizosach
22	41	54.7	153	2 Q9ADL7	Q9ADL7 polyanthum
23	41	54.7	185	10 Q8S7G3	Q8S7G3 oryza sativ
24	41	54.7	216	6 Q9N223	Q9N223 gorilla gor
25	41	54.7	396	10 Q80871	Q80871 arabidopsis
26	41	54.7	542	4 Q9BRG0	Q9BRG0 homo sapien
27	41	54.7	608	3 Q8TFE4	Q8TFE4 gaeumannomy
28	41	54.7	922	10 Q9M497	Q9M497 arabidopsis
29	41	54.7	944	3 P87114	P87114 schizosach
30	41	54.7	948	10 Q9S749	Q9S749 arabidopsis
31	41	54.7	1449	4 Q94837	Q94837 homo sapien
32	41	54.7	2123	10 Q9LPR3	Q9LPR3 arabidopsis
33	40	53.3	91	12 Q8Q258	Q8Q258 kaikauru vi
34	40	53.3	91	12 Q8Q243	Q8Q243 sango virus
35	40	53.3	91	12 Q8QPV6	Q8QPV6 peaton viru
36	40	53.3	91	12 Q8QPV4	Q8QPV4 peaton viru
37	40	53.3	91	12 Q8QHN1	Q8QHN1 peaton viru
38	40	53.3	131	4 Q9P0F9	Q9P0F9 homo sapien
39	40	53.3	131	4 Q96PG9	Q96PG9 homo sapien
40	40	53.3	133	4 Q8WY82	Q8WY82 homo sapien
41	40	53.3	141	11 Q9R0E5	Q9R0E5 marmota mar
42	40	53.3	150	2 Q9F804	Q9F804 erwilia amy
43	40	53.3	247	5 Q62268	Q62268 canorhabdi
44	40	53.3	254	3 Q9P960	Q9P960 aspergillus
45	40	53.3	262	16 Q84111	Q84111 chlamydia t

#### ALIGNMENTS

RESULT 1	Q9SUV8	PRELIMINARY:	PRT:	490 AA.
AC	Q9SUV8	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Hypothetical 53.0 kDa protein.			
GN	F4F15.280.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_TaxID=3702;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RA	Alcaraz J.P., Clabault G., Cottet A., Mache R., Mewes H.W., Lemcke K.,			
RL	Mayer K.F.X., Querier F., Salanoubat M.;			
RL	Submitted (Apr-1999) to the EMBL/Genbank/DBJ databases.			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RA	EU Arabidopsis sequencing project;			
RL	Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: AL049711; CAB41337.1; -			
KW	Hypothetical protein.			
SQ	SEQUENCE 490 AA: 53038 MW: 3784B06F9580447C CRC64;			
Query Match:	65.3%:	Score 49:	DB 10:	Length 490;
Best Local Similarity:	73.3%:	Pred. No. 2.4:		
Matches	11;	Conservative	1;	Mismatches 3; Indels 0; Gaps 0;
QY	1 SSSILCSPPLSLSLV 15			
Db	129 SSSILMDVPPLSLV 143			
RESULT 2				
Q8R4H6	PRELIMINARY:	PRT:	1298 AA.	



Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 SSILCSPPLSLSLV 16  
:||:||||:| 1:  
Db 12 TSILCSPMTSLKTSI 26

RESULT 5  
09MAAO PRELIMINARY; PRT: 490 AA.

AC 09MAAO:  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Putative branched-chain amino acid aminotransferase.  
GN T12H1.16.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,  
RA Rongling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;  
RT "Arabidopsis thaliana chromosome III BAC T12H1 genomic sequence."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AC009177; AF27025.1; -;  
DR HSP: P00510; IASG.  
DR InterPro: IPR001544; Amino-tran\_4.  
DR Pfam: PF01063; aminotran\_4; 1.  
DR ProDom: PD001961; Amino-tran\_4; 1.  
KM Aminotransferase; Transferase.  
SQ SEQUENCE 490 AA; 54829 MW; D29B6876564B36BD CRC64;

Query Match 61.3%; Score 46; DB 10; Length 490;  
Best Local Similarity 68.8%; Pred. No. 7.7;  
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSILCSPPLSLSLV 16  
:||||:||||| 1:  
Db 210 SSILCSPPLSLSLV 225

RESULT 6  
08M027 PRELIMINARY; PRT: 555 AA.

AC 08M027:  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Putative branched-chain amino acid aminotransferase  
DE (AT9G05190/T12H1.16).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_Taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,  
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Soultwick A.,  
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Arabidopsis cDNA clones";  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shin P., Banh J.,  
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
RA Seki M., Soultwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RT "Arabidopsis ORF clones";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF462849; AL158936.1; -;  
DR EMBL: AY090279; AL159040.1; -;  
DR InterPro: IPR001544; Amino-tran\_4.  
DR Pfam: PF01063; aminotran\_4; 1.  
DR ProDom: PD001961; Amino-tran\_4; 1.  
KW Aminotransferase; Transferase.  
SQ SEQUENCE 555 AA; 62212 MW; 670A17B8D74B5D86 CRC64;

Query Match 61.3%; Score 46; DB 10; Length 555;  
Best Local Similarity 68.8%; Pred. No. 8.7;  
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSILCSPPLSLSLV 16  
:||||:||||| 1:  
Db 241 SSILCSPPLSLSLV 256

RESULT 7  
0912A3 PRELIMINARY; PRT: 724 AA.

AC 0912A3:  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Propionyl CoA-carboxylase alpha-subunit.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6;  
RA Miyazaki T., Ohura T., Kobayashi M.;  
RT "Fatal propionic acidemia in mice lacking propionyl-CoA carboxylase  
RT and its rescue by postnatal, liver-specific supplementation via a  
RT transgene";  
RL J. Biol. Chem. 0:0-0(2001).  
DR EMBL: AY046947; AL102364.1; -;  
DR InterPro: IPR001882; Biotin\_attach.  
DR InterPro: IPR000089; Biotin\_11poyl.  
DR InterPro: IPR000901; CPSase.  
DR Pfam: PF02785; Biotin\_carb\_C; 1.  
DR Pfam: PF00364; Biotin\_11poyl; 1.  
DR Pfam: PF00289; CPSase\_L-chain; 1.  
DR Pfam: PF02786; CPSase\_L-D2; 1.  
DR PROSITE: PS00188; BIOTIN; UNKNOWN\_1.  
DR PROSITE: PS00866; CPSASE\_1; UNKNOWN\_1.  
DR PROSITE: PS00867; CPSASE\_2; UNKNOWN\_1.  
SQ SEQUENCE 724 AA; 79660 MW; DB5EC4CBF40739FB CRC64;

Query Match 61.3%; Score 46; DB 11; Length 724;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSILCSPPLSLSLV 16  
:||||:||||| 1:  
Db 654 TSILCSPMPGVVAV 669

RESULT 8  
0922N3 PRELIMINARY; PRT: 724 AA.

ID 0922N3  
AC 0922N3:

```
DR 01-DEC-2001 (TREMblrel. 19, Created)
DR 01-DEC-2001 (TREMblrel. 19, last sequence update)
DR 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Similar to propionyl coenzyme A carboxylase, alpha polypeptide.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC006915; AA06915.1;
DR InterPro: IPR001882; Biotin_attach.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR009301; CPSase.
DR Pfam: PF02785; Biotin_carb_C; 1.
DR Pfam: PF03564; Biotin_lipoyl; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR PROSITE: PS00188; BIOTIN; UNKNOWN_1.
DR PROSITE: PS00866; CPSASE_1; UNKNOWN_1.
DR PROSITE: PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 724 AA; 79940 MW; 8B64661F9E399A2 CRC64;

Query Match
Best Local Similarity 50.0%; Score 46; DB 11; Length 724;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLSV 16
DB 654 TSSILCSMPGVYAV 669

RESULT 9
ID Q15134 PRELIMINARY; PRT; 887 AA.
AC Q15134;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
DE Phosphatidylinositol 3-kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-280 FROM N.A.
RX MEDLINE-95354652; PubMed-7628435;
RA Volinia S., Dhand R., Vanhesebroeck B., MacDougall L.K.,
RA Zvelebil M.J., Domin J., Panaretou C., Waterfield M.D.;
RT "A human phosphatidylinositol 3-kinase complex related to the yeast
RT Vps34p-Vps15p protein sorting system.";
RL EMBL J. 14.3339-3348(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Volinia S.;
RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z46973; CA87094.1;
DR InterPro: IPR000088; C2;
DR InterPro: IPR001263; PI3K.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR000403; PI3_P14_kinase.
DR Pfam: PF00613; PI3K; 1.
DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00454; PI3_P14_kinase; 1.
DR SMART: SM00239; C2; 1.
DR SMART: SM00145; PI3K; 1.
DR SMART: SM00146; PI3K_C2; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS00920; PI3_4_KINASE_3; 1.
```

```
KW Kinase.
SQ SEQUENCE 887 AA; 100988 MW; 2078660E7D9FEF187 CRC64;

Query Match
Best Local Similarity 58.7%; Score 44; DB 4; Length 887;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLS 13
DB 442 SSQITSPSPSVS 454

RESULT 10
ID Q8W6E1 PRELIMINARY; PRT; 82 AA.
AC Q8W6E1;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Putative major coat protein.
OS bacteriophage VSKK.
OC Viruses; ssDNA viruses; Inoviridae.
OX NCBI_TaxID=180503;
RN [1]
RP SEQUENCE FROM N.A.
RA Sinha K.K., Ghosh A.;
RT "Complete nucleotide sequence of single-stranded filamentous
RT vibriophage VSKK.";
RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF452449; AAL40839.1;
SQ SEQUENCE 82 AA; 8804 MW; 62BE0B7157F62D86 CRC64;

Query Match
Best Local Similarity 57.3%; Score 43; DB 9; Length 82;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSSILCSPLPSLS 13
DB 16 SVICPMPMSMS 26

RESULT 11
ID Q9N7D6 PRELIMINARY; PRT; 138 AA.
AC Q9N7D6;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
DE Possible serine/threonine kinase pkn8 (Fragment).
GN LM28.292.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL390935; CAC01001.1;
DR InterPro: IPR000719; Euk_pkinase.
DR PRODOM: PD000001; Euk_pkinase; 1.
KW Kinase.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15163 MW; 2AF74AF488B12D6 CRC64;

Query Match
Best Local Similarity 57.3%; Score 43; DB 5; Length 138;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SSSILCSPLPSLS 15
||||| : |||
```

Db 120 SSSICSPPLPSLSL 134

# RESULT 12

096EH4 PRELIMINARY; PRT; 189 AA.

AC 096EH4; 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE Similar to RIKEN CDNA 4933430F08 gene.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LUNG;  
 RA Strausberg R.;  
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC012330; AAH12330.1; -  
 SQ SEQUENCE 189 AA; 19503 MW; BCID0F5B688D62E7 CRC64;

Query Match 57.3%; Score 43; DB 4; Length 189;  
 Best Local Similarity 71.4%; Pred. No. 10;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SSSICSPPLPSLSL 14  
 DB 60 SPSLQCPPLPSLSL 73

# RESULT 13

09ES77 PRELIMINARY; PRT; 3567 AA.

AC 09ES77; 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Polydom protein precursor.  
 GN POLYDOM.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/HEUSIC;  
 RX MEDLINE=20517255; PubMed=11062057;  
 RA Gilges D., Vinit M.-A., Callebaut I., Coulombel L., Cacheux V.,  
 RA Romeo P.-H., Vigon I.;  
 RT "Polydom : a secreted protein with pentraxin, complement control  
 RT protein, epidermal growth factor and von Willebrand factor A  
 RT domains";  
 RL Biochem. J. 352:49-59(2000).  
 DR EMBL; AF206329; AAG32160.1; -  
 DR HSPD; P00740; LEDM.  
 DR MGD; MGI:1928849; Polydom.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF\_CA.  
 DR InterPro: IPR001438; EGF\_II.  
 DR InterPro: IPR003410; Hyalin.  
 DR InterPro: IPR001759; Pentaxin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR002035; WVF\_A.  
 DR Pfam: PF00008; EGF; 10.  
 DR Pfam: PF02494; HXR; 2.  
 DR Pfam: PF00354; pentaxin; 1.  
 DR Pfam: PF00084; sush1; 33.  
 DR Pfam: PF00092; wvf; 1.  
 DR PRINTS; PR00010; EGFBL00D.

DR PRINTS; PR00895; PENTAXIN.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR Prodom; PD002153; Pentaxin; 1.  
 DR SMART; SM00032; CCP; 34.  
 DR SMART; SM00181; EGF; 15.  
 DR SMART; SM00179; EGF\_CA; 9.  
 DR SMART; SM00001; EGF\_II; 3.  
 DR SMART; SM00159; PTX; 1.  
 DR SMART; SM00327; WVF; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 5.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_9.  
 DR PROSITE; PS01186; EGF\_2; 11.  
 DR PROSITE; PS01187; EGF\_CA; 6.  
 DR PROSITE; PS0234; WVF; 1.  
 KW Signal.  
 FT SIGNAL.  
 SQ SEQUENCE 3567 AA; 387391 MW; 6FBAB276E12293E5 CRC64;

Query Match 57.3%; Score 43; DB 11; Length 3567;  
 Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 SSSICSPPLPSLSV 16  
 DB 1622 SSMFCSDCPDLEGSV 1637

# RESULT 14

090WV3 PRELIMINARY; PRT; 113 AA.

AC 090WV3; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 DE Fas-binding protein Daxx (Fragment).  
 GN DAXX OR DAXX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129SVJ;  
 RA Rowen L., Qin S., Madan A., Loretz C., James R., Dors M., Mix L.,  
 RA Hall J., Lasky S., Hood L.;  
 RT "Sequence of the mouse major histocompatibility locus class II  
 RT region";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF100956; AAC69891.1; -  
 DR MGD; MGI:1197015; Daxx.  
 DR InterPro: IPR005012; Daxx.  
 DR Pfam; PF03344; Daxx; 1.  
 FT NON\_TER  
 SQ SEQUENCE 113 AA; 12086 MW; DAAE6A725336A511 CRC64;

Query Match 56.0%; Score 42; DB 11; Length 113;  
 Best Local Similarity 69.2%; Pred. No. 9.5;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 SSSICSPPLPSLSL 14  
 DB 68 TSSLCSPSPSL 80

# RESULT 15

09YBN2 PRELIMINARY; PRT; 114 AA.

AC 09YBN2; 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Hypothetical protein APE1566.  
 GN APE1566.

OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
OC Desulfurococcaceae; Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K1;  
RX MEDLINE=99310339; PubMed=10382966;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,  
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,  
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,  
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,  
RT \*Complete genome sequence of an aerobic hyper-thermophilic  
RT crenarchaeon, Aeropyrum pernix K1.";  
RL DNA Res. 6:83-101(1999).  
DR EMBL; AP000062; BAA80566.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 114 AA; 11565 MW; 130EE20695FA670 CRC64;

Query Match 56.0%; Score 42; DB 17; Length 114;  
Best Local Similarity 54.5%; Pred. No. 9.6;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 6 CSPPLSLSLV 16  
I:|:|:|:  
DB 28 CTPPSMNLST 38

Search completed: June 25, 2003, 17:17:00  
Job time : 3.93289 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2003, 20:59:10 ; Search time 107 Seconds  
(without alignments)  
4700.466 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgacacactgttcctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/p/ptodata/1/ina/5A-COMB.seq:\*  
2: /cgn2\_6/p/ptodata/1/ina/5B-COMB.seq:\*  
3: /cgn2\_6/p/ptodata/1/ina/6A-COMB.seq:\*  
4: /cgn2\_6/p/ptodata/1/ina/6B-COMB.seq:\*  
5: /cgn2\_6/p/ptodata/1/ina/PCTUS-COMB.seq:\*  
6: /cgn2\_6/p/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	193	11.8	1555	US-08-809-802-11	Sequence 11, Appl
2	187	11.4	4621	US-09-592-891A-13	Sequence 13, Appl
3	182.2	11.1	2310	US-08-417-330A-15	Sequence 15, Appl
4	181.6	11.1	1609	US-09-592-891A-11	Sequence 11, Appl
5	179	10.9	3958	US-08-435-933-5	Sequence 5, Appl
6	179	10.9	3958	PCT-US96-06035-5	Sequence 5, Appl
7	174.2	10.6	1640	US-09-592-891A-12	Sequence 12, Appl
8	168	10.2	1368	US-09-130-339-1	Sequence 1, Appl
9	161.2	9.8	1297	US-09-130-339-3	Sequence 3, Appl
10	157.8	9.6	1657	US-08-554-659-3	Sequence 3, Appl
11	157.8	9.6	1844	US-09-002-361-1	Sequence 1, Appl
12	156.4	9.5	1866	US-08-417-330A-19	Sequence 19, Appl
13	156.2	9.5	1657	US-08-554-659-1	Sequence 1, Appl
14	156.2	9.5	1732	US-08-417-330A-17	Sequence 17, Appl
15	146.6	8.9	669	US-09-002-361-7	Sequence 7, Appl
16	141.8	8.6	1491	US-08-137-614A-3	Sequence 3, Appl
17	141.8	8.6	1491	US-08-768-301-1	Sequence 1, Appl
18	138.6	8.5	1638	US-08-417-330A-13	Sequence 13, Appl
19	137.6	8.4	1519	US-09-002-361-4	Sequence 4, Appl
20	135.6	8.3	2191	US-08-417-330A-11	Sequence 11, Appl
21	135.2	8.0	1052	US-09-592-891A-8	Sequence 8, Appl
22	131.6	8.0	2066	US-08-072-064-7	Sequence 7, Appl
23	130	7.9	2066	US-08-072-064-2	Sequence 2, Appl
24	128.4	7.8	2066	US-08-072-064-3	Sequence 3, Appl
25	128.4	7.8	2066	US-08-072-064-5	Sequence 5, Appl
26	125.8	7.7	1650	US-08-459-100A-1	Sequence 1, Appl
27	125.8	7.7	1650	PCT-US94-09589-1	Sequence 1, Appl

28	124	7.6	1707	US-08-809-802-7	Sequence 7, Appl
29	119.6	7.3	1970	US-08-072-064-9	Sequence 9, Appl
30	119.6	7.3	1970	PCT-US92-08558-3	Sequence 3, Appl
31	119.6	7.3	2059	PCT-US92-08558-2	Sequence 2, Appl
32	111.2	6.8	1542	US-08-249-112-1	Sequence 1, Appl
33	111.2	6.8	1542	PCT-US95-06556-1	Sequence 1, Appl
34	109	6.6	2635	US-08-768-301-3	Sequence 3, Appl
35	106.2	6.5	2508	US-09-627-650B-8	Sequence 8, Appl
36	106.2	6.5	2508	US-09-436-063C-8	Sequence 8, Appl
37	106.2	6.5	2544	US-09-627-650B-4	Sequence 4, Appl
38	106.2	6.5	2544	US-09-436-063C-4	Sequence 4, Appl
39	106.2	6.5	2601	US-09-627-650B-10	Sequence 10, Appl
40	106.2	6.5	2601	US-09-436-063C-10	Sequence 10, Appl
41	102.8	6.3	762	US-09-592-891A-7	Sequence 7, Appl
42	101.6	6.2	1917	US-09-627-650B-6	Sequence 6, Appl
43	101.6	6.2	1917	US-09-436-063C-6	Sequence 6, Appl
44	97.8	6.0	1479	US-08-249-112-2	Sequence 2, Appl
45	97.8	6.0	1479	PCT-US95-06556-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-809-802-11  
Sequence 11, Application US/08809802  
Patent No. 6455276  
GENERAL INFORMATION:  
APPLICANT: Le Bourdellies, Beatrice  
APPLICANT: Whiting, Paul John  
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT  
TITLE OF INVENTION: OF THE GABA-A RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809, 802  
FILING DATE: 19-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/02323  
FILING DATE: 29-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Yablonsky, Michael D  
REGISTRATION NUMBER: 40,407  
REFERENCE/DOCKET NUMBER: T1292  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-4678  
TELEFAX: 732-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 47...1402  
OTHER INFORMATION:  
US-08-809-802-11

Query Match 11.8%; Score 193; DB 4; Length 1555;  
 Best Local Similarity 53.3%; Pred. No. 6.3e-47;  
 Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

142 CTGAGCAAACTTATGGGGGACATCTGATATGATGACGATGAGCCCAATTTTAA 201  
 167 CTCCCAACCTGGACGGGCTGATAGCCGGTTAGCCCGCACTTCGGCTGGCATCGGA 226  
 202 GGGCCACCGTGAACCTGACCTGCAACATCTTCATCAACGTTTCAGCTCCGACCAAG 261  
 227 GGGCCCCCGTGAATGAGGCCCTTGGCCCTGAGAGTGGCCATGACACCATCTCAGAG 286  
 262 ACCACATGACATACCGGGGATGCTCTTCTGGCGCAACAGTGAAGATGCCACGCGCTG 321  
 287 GCCAACATGAGATACACATGACGAGGTCTCTGACAGCTGGGGGAGACAGAGCTC 346  
 322 TCCATCCGATATATCTGATGATCTCTGACCTGATCCCTCATGCTGATCTATC 381  
 347 TCCATC---AACACACCAAGAGACCTTGGCTGACAGCCGCTTCGAGCAAGCTG 403  
 382 TGGAAAGCCAGACCTCTTCTTGTGATATGAGAAAGGGCCAACTTCATGAGTGACCAAG 441  
 404 TGGGTGCGCGACACCTTCATCTGTAAGCCAGTGGCTGTGTCACGAGCTGACGGTG 463  
 442 GACACAAAGTTACTGCGCATCTTCAAGATGGGAATGCTGTACAGCATCAGGCTGACG 501  
 464 GAGAACAGCTCATCCCGTGCACAGCCGAGGGGTGATCTGTACACATCCGATACCC 523  
 502 CTCATTTTGTCTCTGATGATGACATCAAGAACTTCCCATGACATCAGACGTGACG 561  
 524 TCCACTGTGGCTGCGACATGACATGACCTGGCCAAATTTCCCATGACGAGAGATGCA 583  
 562 ATGACGTTGAGAGCTATCATCTACTGACAGCCCTGCCATCTGTCTACTTCAAGTT 621  
 584 CTGACCTGAGAGACT-----AC 601  
 622 GGCTACACCATGAAAGACCTGTTTGAAGGCTGGAAGATGCTCTGCTGTCACAGTG 681  
 602 GGTACTCATGAGAGACATGCTACTACTGTGTCGAGAGACCAAGACATCAGAGGG 661  
 682 GCTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGGGGATGAGAAAGATCTAGGCTGTT 741  
 662 CTGGAACAGCTGACAGTGGCGCAGTTCCACATCACACGCTACCCCTTCCACAGGAGCTG 721  
 742 ACCAAGCCTACACACAGGGAATTCACCTGCACTCGAGTAAAGTTTCACTGGAACGG 801  
 722 ATGAACCTCAAGTCCGCTGGCCAGTTCCACGCGCTCAGCTGCACTTCCACCTGCGAGG 781  
 802 CAGATGGGCTACTATCTGATTCAGATACATCCCAAGCTACTACTGATCTGATCTGTC 861  
 782 AACCGCGCGTGTACATCATCAATCTACATGCTCCGCTCCGCTGCTGTCGTCATGTC 841  
 862 TGGGTCTCTCTGATCAACATGAGATGCTGCCCCCTGCGCTGAGGCTGGGCAATCAC 921  
 842 TGGGTCTCTCTGATCAGCCAGGCGGGGCGCCGAGGGTCTCTCTAGAGCATCAC 901  
 922 ACCGTGCTACCATACCAACCAAGAGCTGTGGCTCCCGGGCTCTTGGCTTAAAGTGTCC 981  
 902 ACGGTGCTGACATGACAGCTCATGTGTCATGTCCTCTCCCTGACAGGGCATCA 961  
 982 TACGTGAAGCAATGATGATGAGTGGTGTGCTGCTTGTGCTGCTGCTGCTGCTGCT 1041  
 962 GCCATCAAGGACATGACATCTACTTCTGATCTGATCTGATCTGATCTGATCTGATCT 1021  
 1042 CTGAGATATGC 1052  
 1022 GTGAGATACGC 1032

RESULT 2  
 US-09-592-891A-13  
 ; Sequence 13, Application US/09592891A  
 ; Patent No. 6329174

GENERAL INFORMATION:  
 APPLICANT: Xiao-Zhou Michelle Wang  
 APPLICANT: Xavier Georges Sarda  
 APPLICANT: Michael David Tomalski  
 APPLICANT: Vincent Paul Mary Wingle  
 TITLE OF INVENTION: Helicobacter Receptor  
 FILE REFERENCE: A32815 072667.0118  
 CURRENT APPLICATION NUMBER: US/09/592,891A  
 NUMBER OF SEQ ID NOS: 14  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 13  
 TYPE: DNA  
 ORGANISM: Helicobacter  
 FEATURE:  
 NAME/KEY: unsure  
 LOCATION: (764)...(764)  
 OTHER INFORMATION: n is a, t, g, or c  
 US-09-592-891A-13

Query Match 11.4%; Score 187; DB 4; Length 4621;  
 Best Local Similarity 54.3%; Pred. No. 6.9e-45;  
 Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;

208 CCCGTGAACGTGACCTGACATCTTCATCAGATTTTCAAGTTCCTCCGTCACCAAGACCACA 267  
 330 CCAGCGTATGACCGCTCATATATTTGTCGAAGTATTCAAAGATTCATGACGTACCA 389  
 268 ATGAGCTACCGGCTGAATGCTTCTTCCGCAACAGTGAATGATACCAACCGCTGCTAC 327  
 390 ATGGAATACCTCCGTCATGATTAAGCTTCCGCAACAGTGAATGATTAAGTGAACGCTCA 449  
 328 CGAGATATCTGATGATCTCTGACCTGATCCCTTCATATGCTGATCTATCTGAGAG 387  
 450 AATAATCTGAGAGTCCGCTCAATATCTGACATGACCGAAGCAAGCATGCTGATG 509  
 388 CCAGACCTCTTCTTGTATGAGAAAGGGCCAACTTCCATGAGTGAACAGGACCAAC 447  
 510 CCTGATCTATTTCTTCCCAAGAGAAAGGATCTATTTCCCAACATCATCATGCTCCGAC 569  
 448 AAGTTACTGCGCATCTTCAGAAATGGAATGCTGTACAGCAATCAGAGCTGACCTGAT 507  
 570 GTGATACCTCGGATCTTCCCAAGGCAAGCTGTGACAGATCCGAATCTCTTACAG 629  
 508 TTGTCCTGCTGATGAGACCTCAAGAACTTCCCATGACATCCAGAGTGCACAGATCAG 567  
 630 CTCGCTGCGCCATGACCTCAAGTTGATACCCCTGATAGCAAGACCTGCTGCTCAGG 689  
 568 CTGAGACCTCATCATCTGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627  
 690 ATGGCTACT-----ATGGTTGG 707  
 628 ACCATGAAGACCTGCTTGTGAGTGTGGAAGATGCTGCTGCTGCTGCTGCTGCTGCTG 687  
 708 ACCACAGAGACCTGATGTTCTTATGAGAAAGAGGCGCC---CGGTGAGAGTGTGAGAN 764  
 688 GGGGTGACTGCGCCCATTTATCTTGGGAGTGAAGAGATCTAGAGCTGTTGATCAAG 747  
 765 AACTTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821  
 748 CACTACACACAGGAAATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 807  
 822 AAGACTAATACCGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881  
 808 GGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 867  
 882 AGTTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941  
 868 TCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 927  
 942 TCTTCTGCTGACAGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001



;; CURRENT APPLICATION NUMBER: US/09/592,891A  
;; CURRENT FILING DATE: 2001-01-22  
;; NUMBER OF SEQ ID NOS: 14  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 11  
;; LENGTH: 1609  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: PCR amplified fragment of Heliothis DNA cloned  
;; OTHER INFORMATION: Into pCR2.1-TOPO vector (Invitrogen)  
US-09-592-891A-11

Query Match 11.1%; Score 181.6; DB 4; Length 1609;  
Best Local Similarity 54.2%; Pred. No. 1.5e-43;  
Matches 469; Conservative 0; Mismatches 349; Indels 48; Gaps 3;

```

OY 208 CCCGTGACGCTGACCTGCAACATCTTCATACAGTTTCAGCTCCGTCACCAAGACCACA 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 CCAGCGGTAGTGGCGGTGATATTTGTCGAAATATATCAAGATGATGACGTGACACA 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 268 ATGAGTACCGGGTGAATGTCCTGCGGCAACAGTGAATGACCCAGCGCTGCTCTAC 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 ATGGAATACCTCCGTACATTAACGTTTCGGGAACAATGTTAGTGAACGGCTCAATTC 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 328 CGAATATTCCTGATGACCTCTGACCTGATCCCTCATGCTGACCTATCTGAGAG 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 454 AATAATCTGGAGTGCCCTCAAAATACCTGACACTGACTGAAGCAACAGAGCTGATG 513
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 388 CCAGACCTCTTTTGTGTAATGAGAAAGGGCCCACTTCATGAGGTGACCGAGCAC 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 CCGATCTATCTCTCCCAAGAGAGAGAGGATCTTCACAAATCATCATGCGCCGAC 573
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 448 AAGTACTGGCATCTTCAGAAATGGGATGCTGTGACAGTCAAGCTGACCTGACCTCAT 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 574 GTGACATCCGATCTTCTCCCAAGCGAAGCTGCTGTACAGCATCCGAATCTCCCTAGC 633
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 508 TTGCTCTGCTGATGACCTCAAGAACTCCCATGAGATCCAGAGCTGACAGAGTACG 567
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 634 CTCGCGTCCCATGAACTCAAGTTCTACCCCTGATTAAGACAGACTCTGCTCAGG 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 568 CTGAGAGCTCATCCATACCTGAGAGCCCTGCGCATCTGTGACATTTAGTTGGTAC 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 694 ATGGCTAGT-----ATGGTGG 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 628 ACCATGAAGACCTCGTGTGAGTGCTGGAAGATGCTCTGCTGTCAAGTGGCTGAG 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 712 ACCACAGACGATAGTTCATGGAAGAGAGCGAC--CGTGCAGGTGGTGA 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 688 GGGCTGACTGCCCCAGTTTATCTTGGGGATGAGAGATCTAGGCTGTTTACCAG 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 769 AACTTACCTGCTCGGTTAC--GCTGAGAGAGTCTCTCACTGCTGCAACAGT 825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 748 CACTACAAACAGGAGAAATTCACCTGATGAGGTAAAGTTTCAACGAGAGGCAATG 807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 826 AAGACTAATACCGGTAAATACAGTTGCTTAAGGTAACTGCTTTCAACGAGATTC 885
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 808 GGGTACTATCTGATGATGATGATCCCAAGCTTCACTCATGTCATCTGCTGAGT 867
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 886 AGTTACTACTGATCCAGATCTAATCTCGTGCATGCTGTGATCATCGTCTGGGT 945
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 868 TCCTTCTGATCAACATGATGCTGCCCTGCCGCTGGGCTGGGCTGACACCGGTG 927
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 946 TCCTTCTGCTGAGACGAGGAGCTGTGCTGCGAGGGGTCTACATAGAGAGTGA 1005
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 928 CTGACATGACCAACAGAGCTCTGCTCCGGGCTCTTTGCTTAAGGTGCTGACGTG 987
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1006 CTTACATATGGCAGCGCTGCTGAGGATCAACGCGCTTACACCGGCTCTTACAGC 1065
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 988 AAGCAATGACATCTGATGCTGTGCTGTCTTTGTTGTTGCTGCTGCTGCTGAG 1047
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1066 AAGGCATGATGCTGAGACTGGGTATGTCTACATTTCTGAGGCGCTACTAGAG 1125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

OY 1048 TATGCTGCATTAATTTTGTCTCG 1073
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1126 TTTGCGCTGCTCAACTATGCGTCTCG 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 5

```

US-08-435-933-5
; Sequence 5, Application US/08435933
; Patent No. 5693492
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Arena, Joseph P.
; APPLICANT: Parens, Phillip S.
; APPLICANT: Liu, Ken K.
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
; TITLE OF INVENTION: CHANNELS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,933
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen, III John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-435-933-5
```

Query Match 10.9%; Score 179; DB 1; Length 3958;  
Best Local Similarity 53.7%; Pred. No. 1.4e-42;  
Matches 473; Conservative 0; Mismatches 360; Indels 48; Gaps 3;

```

OY 193 AATTTTAAAGCCGCCACCGGTGACGTGACATCTTCAACAGTTTCACCTCC 252
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 570 AATGGACACAGATGTCGCCGATGATGATTAATCTGTCGCAATATTAAGAG 629
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 253 GTACCAAGACCAATGACTACCGGGTGAATGTCCTTTGGGCAACAGTGAATGAC 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 630 ATTAGTATATTTAAATGAGTACAGTGTCACTTAACCTTCCTGAACAGTGA 689
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 313 CCAGCGCTGTCTACCGGAATATCTGATGACTCTGTGACCTGATCCCTCATGCTG 372
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 690 GAACGCTCAAGTTTGCAGATATCCAGGGTGGCTTAAGTATCTGACCTGAGGAG 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 373 GACTATCTGGAACCAACCTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 750 AACCGCTGTGAGATGGCCGATCTTTCTCTCCAAAGAGAGGAGACACTTCCACA 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 433 GTGACACGAGCAACAAGTACTGAGGATTTCAAGAAAGGAAAGTGGTACACATC 492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 810 ATCATCATGCCCAATGATGATATTCGATCTTCCCAACAGGATGTGCTATATTA 869
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

OY	493	AGGCTGACCCCTCATTTTGTCTCCGTGATGAGACCTCAAGAACTTCCATGGACATCCAG	552
Db	870	CGTATCTCGCTGCAATTTGGCTCGCCAAATGAACCTAAAGCTGTATCCGCTGGATGAGTAG	929
OY	553	ACGTGCAAGATGACAGCTTGAGAGCTCATCCATACCTGTGACACCCCTCTGCCATCTGTCA	612
Db	930	ATTCGTCTACATACGAGATGGCCAGCT-----	954
OY	613	CTTTTCAGTTGGCTACACCATTAAGAAAGCTCGTGTTTGATGAGCTTGAAGATGCTCGCT	672
Db	955	-----ATGGCTGTGACACCAACAGACTTGGTCTTCCTGTGGAGAGAGGGGATC---CC	1004
OY	673	GTCCAAAGGCTGAGGGGCTGACGTCTGGCCCGAGTTTATCTTGGGGGATGAGAAGATCTA	732
Db	1005	GTACAGGTGGTAAAGAACTTACACCTTACTCTGCTTCACTAGAGAAATTTTCTGACT---	1061
OY	733	GGCTGTGTATCCAAGCACTACACACAGAGGAAATTCACCTCATCGATGAGTAAAGTTTAC	792
Db	1062	GATTACGTGTACAGTAAACCAACACCGGTGAATACACTTGGCTCTAAAGTGGATTTACTA	1122
OY	793	CTGGAACGGCAGATGGGCTACTATCTGATTCAGATGTACATGCCAGCTACTATCGTC	852
Db	1122	TTACAGGGAGATTTCTCATATTACTATTAAACAAATTTTATACATGCTGTATGTTGTC	1181
OY	853	ATCCTGTCTTGGGTCTCCTTCTGGATCAACATGATGCTGCCCTCGCCGTGTGGGCTG	912
Db	1182	ATTGTATTCATGGGTATCTATTTCTGGCTGGATCAAGGAGAGTATACCGCCGAGGTGCACCTG	1241
OY	913	GGCATCACACCGGTCTCACCATGACCAACCCAGAGCTGTGGCTCCGGGCGCTTTGGCT	972
Db	1242	GGTGTCAACACCCCTGCTGACCATGATGCCACCCAGACGTGGGCATTAACGCCCTCCCTGCG	1301
OY	973	AAGGTGTCCTACGTGAAGCAATCGACATCTGGATGAGCTGTGTGCTCTCTTTGTGTTT	1033
Db	1302	CCCGTTTCTTACGAAGGCGCATGTGATGTGGACAGGCGTGTGCTGACGTTGCTGTTC	1361
OY	1033	GCTGCTTGTGCTGGAGTATGCTGCCATTAATTTTGTTCG	1073
Db	1362	GGGGCCCTGTGAGTTCGCCCTGGTGAACATATGATCCG	1402

RESULT 6  
PCT-US96-06035-5  
Sequence 5, Application PC/TUS9606035  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Arena, Joseph P.  
APPLICANT: Parese, Philip S.  
APPLICANT: Liu, Ken K.  
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jody M. Glesser  
STREET: 126 East Lincoln Avenue - P.O. BOX 2000-0907  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06035  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Glesser, Jody M.  
REGISTRATION NUMBER: 32,838  
REFERENCE/DOCKET NUMBER: 19264 PCT  
TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: (908) 594-3046
: TELEFAX: (908) 594-4720
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3958 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
: PCT-US96-06035-5.

```

Query Match	Best Local Similarity	10.9%	Score 179:	DB 5:	Length 3958:
Matches	Conservative	0:	Mismatches 360:	Indels	Gaps 3
QY	193	AATTTTAAAGCCCAACCCGCTGACGCTGACCAACATCTTCATCAACAGTTTCAGCTCC	252		
Db	570	AAATGGACAGATGATGTCGCCGATAGTCAAGATCAATCTATCTGTTGGCAGTATTTAGAC	629		
QY	253	GTCACCAAGACCCACATATGGACTACCGGGTGAATGTCTTTTCGGCAACAGTGGATGAC	312		
Db	630	ATTAGTGTATTTAAATGAGTACAGTGTGACGTTAACCTTCGCCGAACAGTGCAGGAT	689		
QY	313	CCAGCGCTGCTACCCGGAATATCCCTATGACTCTCTGAGCCTGATATCCCTCATGGCTG	372		
Db	690	GAGCGCTCAAGCTTGACGATATCCAGGCTGCCCTAAAGATATCTGACCTGACGGAGCG	749		
QY	373	GACTCTATCTGGAAGCCAGACCTCTCTTCTGCTAATGAGAAAGGGCCCACTTCCATGAG	432		
Db	750	AACGGCGTGGATGTCGCCGATCTTTTCTTCTGGAACGAAAGGAGGACACTTCCACAAAC	809		
QY	433	GTCACACGCAACAAAGTTACTGGCATCTTTCAGAAATGGGAATGTCTCTGACGACATC	492		
Db	810	ATCATCAACGCCCAATGTGTATATTCGCAATCTTCCCAACGGGATCTGTCTATATAGTATA	869		
QY	493	AGGCGTACCTCATTTTGTCTGGCTGTATGGACCTCAAGAACTTCCCATGACATCTCAG	552		
Db	870	CGTATCTGCTGACATTTGGCTGCCCAATGAACCTTAAAGCTGTATCCGCTGGATGAGAC	929		
QY	553	ACGTGCAGATGACAGCTTGAGAGCTCATATCTCTGACGCCCTCTGCATCTCTGTCA	612		
Db	930	ATCTGCTCACTACGATGGCCAGCT-----	954		
QY	613	CTTTCAGTTGGCTACACCATGAAGACCTCGTGTGATGGCTGGAGATGCTCTCTCT	672		
Db	955	-----ATGCTGTGACCAACCAACAGCACTGTGCTTCTGCTGGAAAGAGGGCGATC---CC	1004		
QY	673	GTTCAAGGTGGTGAGGGGCTGACTCTGCCCCAGTTTATCTTGGCGGATGAGAAAGATCTA	732		
Db	1005	GTAACAGTGGTAAGAACTTACCTACTCTGCTTACACACTGAGAAAGTTCTTGACT---	1061		
QY	733	GGCTGTGTACCAAGCACTACAACACAGGAAATTCACCTGCATCGAGGTAAAGTTTCAAC	792		
Db	1062	GATTACTCTACAGTAAGAAACCAACACCGGTGAATACAGTGGCTCAAAAGTCATCTACTA	1121		
QY	793	CTGGAAGCGCAGATGGGCTACTATGTGTGATTCAGATGACATCTCCAGCTACTCATCTGC	852		
Db	1122	TTTCAGGGCAGAAATCTTCATATTACTTAATACAAATTTTATATACCATGCTGATTTGTTGGC	1181		
QY	853	ATCTGTCCTGGGTCCTCTCTGATCAACATGATGTGCCCCCTGCCGCTGGGGCTG	912		
Db	1182	ATTGTATCATGGTATCTATTCTGGCTGGATCAAGAGACATGACGGCGGATGTCACTG	1241		
QY	913	GGCATCACCACTGCTCACCATGACCAACCCAGAGCTGTGGCTCCGGGCTCTTGTGCT	972		
Db	1242	GGTGCACACCTGCTGACCATGGCCACCCAGAGTGGGGCATTAAGCGCTCCCTGCGG	1301		
QY	973	AAGGTGCTCTACGTAAGCAATGACATCTGATGGCTGTGTCTCTTGTGTTTC	1033		
Db	1302	CCCGTTCTATACCAAGGCCATGATGTGTGACAGCGTGTGTCTGACGTTCTGTGTC	1361		
QY	1033	GCTGCTTGTGAGATGCTGCTCATTAATTTTGTTCGC	1073		



D	b		573	-----TGGTTGAGCACAACAAAGATCTGGT	596
O	y		645	GTTTGAA TGCGTCGGA AATGCTCTCTGCTGTCCAA GTGGCGGTGACTCTGCCCA	704
D	b		597	GT TTTTGTGGA AGAAGGTGACC CG--GTGCAGGTTGTCAAGA ATCTACATCTGCCAG	653
O	y		705	GTTTATCTGGGGATGAGAAGATCTTAGGCTGTTGACCAAGACATA CAACACAGGGA	764
D	b		654	GTTTACGTT---GGAGAA GTTCTTGACGGA TTATTGTACAGCAA ACCAATACCGGTGA	710
O	y		765	ATTCCACCTGCATCGAGGTAAAGTTTCACTGGAACGCGAGATGGGCTACTTGTATTC	824
D	b		711	ATACAGTTCCTGGAAGTTCACCTGCTCTTTAAACGAGA GTTCTCGTACTACCTGATCCA	770
O	y		825	GATGTACATCCCCAGCCTACTCATGTGATCTGTCCTGGTCTCTTCTTGATCAACAT	884
D	b		771	GATTCATATTCCTCTGTTGCATGTTGGATGATCGTTCCCTGGGCTGCTGTGTTGAGCA	830
O	y		885	GGATGCTGCCCTCCCGGTGCGGCGCTGGGCAATACCACCGTGCATCACATGACACCCA	944
D	b		831	GGGAGCGGTTCCGGCCAGAGATCATCGGGGTGATACCACATCTCTCA CCAATGCGCACCCA	890
O	y		945	GAGCTGTGCTCCCGGGCCTCTTTTGCCTAAGGTGTCCTACGTGAAGCAATCGACATCTG	1004
D	b		891	GAGCTGGGGCATAAAGCCCTCCCTCGCCGCAAGTGTCTCA CCAAAAGCATCGACGTCTG	950
O	y		1005	GATGCTGTGTGTCTGCTCTTTGTGTGCGCTGCTTGCTGGAGTATGCTGCCATPAATTT	1064
D	b		951	GACCGGAGTCTGCCCTCACGTTCTGCTTGGGGCTTGTCTCGAATTCGCCCTCGTCACTA	1010
O	y		1065	TGTTTC 1070	
D	b		1011	CGCCTC 1016	

RESULT 9  
US-09-130-339-3  
Sequence 3, Application US/09130339  
Patent No. 6358701  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Pareess, Philip S.  
APPLICANT: Warneke, Jeffrey W.  
APPLICANT: Etlter, Adrian  
APPLICANT: Cohen, Charles J.  
TITLE OF INVENTION: DNA MOLECULES ENCODING CTENOCEPHALIDES  
FILE REFERENCE: 20029  
CURRENT APPLICATION NUMBER: US/09/130.339  
CURRENT FILING DATE: 1998-08-06  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 1297  
TYPE: DNA  
ORGANISM: ctenocephalides felis  
US-09-130-339-3

Query Match	9.8%	Score 161.2	DB 4	Length 1297
Best Local Similarity	53.9%	Pred. No. 1.4e-37		
Matches	433	Conservative	0	Mismatches 322; Indels 48; Gaps 3

  

Oy	267	AATGACTACCGGGTGAATGCTTCTTGGGGCAACAGTGAATGACCCAGCCCTGTCTCTTA	326
Db	190	ACTGGAAATACAGTGTCCAGTTAACTTTCACGGAAATGCGAGATGAGAGTTGAATTT	249
Oy	327	CCGGAATATATCTATGACTCTCTGAGACCTCGATGCCCAATGCTGACACTATCTGGAA	386
Db	250	TAAAGACTTTGAGAGTCTCTTTAAATTAATCTTAACACTAACCAGCAAGTCTGTATGAT	309
Oy	387	GCCAGACCTCTTTCTTGTCTAATGAGAAAGGGCCAACTTCATGAGGTGACACCGACAA	446

Db	310	GCCGCAATTGTTCTTTGGCAATGAAAGAGGGCCACTTTCACAAATCATCATGCCGA	369
OY	447	CAAGTTACTGGGCATCTTCAAGAATGGGAATGTGCTGTACAGCATCAGGCTGACCCAT	506
Db	370	CGTCTACATTCGTATTTTTCCTTACGGTTCCGTACTATACAGCATCAGGATATCCCTTAC	429
OY	507	TTTTCCTGCCCTGATATGGACATCTCAAGAAATTTCCCCATGGAACATCCAGAGGTGCAGATCA	566
Db	430	TTTGGCGTCTCTATGAACTCGAAACTATTCGGCTCGATAGGCAAGGTGCTTCTCCG	489
OY	567	GCTTGAGAGCTCATTCACATACCTCGACGCCCTCGCCATCTGTACATTTGAGTGGCTA	626
Db	490	GATGGCCAGTTA-----TGGTTG	507
OY	627	CACCATGAAACCTCGTGTTTGATGGGCTGGAAATGCTCCTGCTGTCCAGTGGCTGA	686
Db	508	GACCAACAACCATCTGGTGTTTTGTGGAAAGGAAGTGACCCG--GTGACAGTTGTCAA	564
OY	687	GGGGCTGACTGTGCCCGAGTTTATCTTCGGGATAGAGAAGATCTAGGCTGTGTACCA	746
Db	565	GAATCTACATCTGCCACAGTTTACGTT--GGAGAAGTTCTTGAGGAAATTATTGTAAACG	621
OY	747	GCACCTACACACAGGAAATTCACCTGCATCGAGGTAAGTTTCACTCGGAACGGCAGAT	806
Db	622	CAAAACCAATACCGGTGAATACAGTTGGCTGAAGGTCGACCGCTCTTTAAACGAGAGTT	681
OY	807	GGGGTACTATCTGATTGATGTACATCCCCAGCCCTACATCATCGTATCCCTGCGGGT	866
Db	682	CTCGTACTACTGATCAGATCTACATTCCTTTGTGCATGTTGGATGATGTTTCTCGGGT	741
OY	867	CTCCTTCTGATCAACATGATGCTGCCCTGCCCTGCCGTGGGGCTGGGCATCACACCGT	926
Db	742	GTCGTTCTGGTGGACCAAGGAGCGCGTTCCGGCCAGAGATATCACTGAGGTGTACCACTCT	801
OY	927	GCTCACCATGACACCCAGAGCTGTGGCTCCGGGCCCTTTGGCTTAAGGTGTCACAGT	986
Db	802	CCTCACCATGGCCACCCAGACGTCGGGCATTAACGCTTCCTCGCGCCACAGTCTCTACAC	861
OY	987	GAAGGCAATGACATCTGGATGGCTGTGTGCTCTTTGTGTTCCGTGCCCTTGCTGGA	1041
Db	862	AAAAGCCATGACGCTGGACCGGAGTGTGCTCAAGTTCGTTCTCGGGGCTTGTCTCGA	921
OY	1047	GTAATGCTGCCAATAATTTTGTTC	1070
Db	922	ATTGCGCCCTGTCACATCGCCTC	945

RESULT 10  
 US-08-554-659-3  
 : Sequence 3, Application US/08554659  
 : Patent No. 5767261  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Wingate, Vincent  
 : APPLICANT: Wolff, Mark  
 : TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL  
 : TITLE OF INVENTION: AND METHODS OF USE THEREOF  
 : NUMBER OF SEQUENCES: 12  
 :  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Scully, Murphy & Presser  
 : STREET: 400 Garden City Plaza  
 : CITY: Garden City  
 : STATE: New York  
 : COUNTRY: USA  
 : ZIP: 11530-0299  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: PatentIn Release #1.0, Version #1.25  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/554,659  
 :  
 : FILING DATE:  
 : CLASSIFICATION: 436

ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9732  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1657 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 115..1581  
US-08-554-659-3

Query Match 9.6%; Score 157.8; DB 1; Length 1657;  
Best Local Similarity 51.6%; Pred. No. 1.7e-36;

Matches 474; Conservative 0; Mismatches 397; Indels 48; Gaps 3;

```

QY 137 ATTCTAGACAACTTATGGGCGCAATCTGGATATGATGCCAGATTGCGCCCAATT 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 239 ATATCTACACCTTTTGGATTGCTAAGTAAAGTACGACAAAGAGTGAAGCGCAACT 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 TTAAGGGCCCGCGAAGCGTACCTGACATCTTCATCAAGTTTCACTCCGCTCA 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 299 ATGGAGAGCCCGCATGTGAGTGGAGATGACATGATGCTCTCCATCAGCTCTTAT 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 257 CCAAGACCAATGATACCGGGTAAATGCTTCTTGGCGCAACAGTGAACCCAC 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 359 CTGAGTGAATGATTTTACCCGATGATTTCTACTGACAAATTTTGGACAGACCCA 418
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 317 GCCTGCTCCACGAAATTC---TGATGACTCTGTGACCTCCATCCCTCATGCTG 373
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 419 GCTTGTCTTCAAAAAGACGCGGTGTGAGACTGTGCTCGCTCGCTCGAATTTATTA 478
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 374 ACTTATCTGGAAGCAGACTCTTCTTGTAAATGAGAAAGGGCCACTTCATGAGG 433
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 479 GAAACATATGGGTACCCGACACTTCTTTGTTAAGCAAAAAGATCATATTTCCACAT 538
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 434 TGAACACGACAAAGTACTGCGCATCTTCAAGATGGAGTGTCTGTACAGCATCA 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 539 CTACTACAGCAAGCATCATACGATCATCATCTGATCTATCTAGAGATATAA 598
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 GGCTGACCTCATTTTGTCTGCTGATGAGACCTCAAGAACTCCCGATGACATCCAA 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 599 GACTGACTATACCCGCTTCTTGTCCGATGATTTGCAAGTATTTTCCGATGACCGTCAAT 658
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 CGTGACAGTGCAGCTTGAAGCTCATCATCTGACAGCCCTCGCATCTGTCTAC 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 659 TATGCAATTTGAATCGAAAGTT----- 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 614 TTTCAAGTTGGCTACACCATGAAAGCTCTGTTGATGGCTGGAAGATGCTCTGCTG 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 683 -----TTGGCTACACCATGCGGACATCCGATACAAAGATGAGAGGGGCCCACTGAG 736
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 674 TCCAGTGGCTAGGGGCTGACTTCCCGCATTTATCTTGGGGATGAGAGATCTGAG 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 737 TGGGTGTGCGAGCAAGTGTCTTGGCCCAATTTCAAGTGTGGGCCACCGGACGGGG 796
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 734 GCTGTTGTACAGCACTACACAGAGGAATTCACCTGCTGCTGAGTAAAGTTTCCACC 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 797 CCATGAGAGTTTCTC---TTAGACACAGGAACTACTCTGCTGCGATGTGAATTTCAAT 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 794 TGGAAAGCGAGATGGGCTACTATCTGATTCAGATGATACATCCCGACCTACTCATCTGTA 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 854 TTGTAGGCTCGATGGATCTACTATTAATTCAGATTTATATTCGCTGCGCTAATTTGTCA 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 854 TCTGTCTCGGCTCTCTTCTGTGATCAACATGATGATGCTCCCTGCGCTGTGGGCTG 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

DB 914 TTATATCTTGGGTATCATTTTGGTGAATGCAAAATGCGACACTGCAAGGATGACATGAG 973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 914 GCATACACACCTGCTCACCATGACACCCAGCAGCTGTCCCGGGGCTTTGGCTTA 973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 974 GTGTACAACTGTATTGACGACGACGCTCATGTCTGTCACGAATGGGCTGCGCCA 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 974 AGTGTCTACGTGAAGGCATGACATCTGATGATGCTGTGCTCTTGTGTTGCTG 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1034 AGATCTCATATGTCAAGTCAATGATGATGATCTATCTGGAACCTGTTTGTGATGCTTCA 1093
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1034 CTGCTTGTGATGATGCT 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1094 CCAGTTACTAGAAATATGC 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 11

US-09-002-361-1

Sequence 1, Application US/09002361

Patent No. 6329516

GENERAL INFORMATION:

APPLICANT: Halling, Blak

TITLE OF INVENTION: Lepidopteran GABA-gated Chloride

TITLE OF INVENTION: Channels

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Dechert Price & Rhoads

STREET: 997 Lenox Drive, Building 3, Suite 210

CITY: Lawrenceville

STATE: NO

COUNTRY: USA

ZIP: 08543

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/002,361

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Bloom, Allen

REGISTRATION NUMBER: 29,135

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 609-520-3214

TELEFAX: 609-520-3259

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1844 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 104...1591

OTHER INFORMATION:

US-09-002-361-1

Query Match 9.6%; Score 157.8; DB 4; Length 1844;

Best Local Similarity 51.6%; Pred. No. 1.7e-36;

Matches 474; Conservative 0; Mismatches 397; Indels 48; Gaps 3;

```

QY 137 ATTCTAGACAACTTATGGGCGCAACATCTGATATGATGCCAGATTGCGCCCAATT 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 ATATCTACACCACTTTTGGATTGCTAAGTAAAGTACGACAAAGAGTGAAGCGCAACT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 197 TTAAGGGCCCGCGAAGCGTACCTGACATCTTCAATCAACAGTTTCACTCCGCTCA 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



Db 288 ATGAGAGCCCGCAGTGATGTGGAGTCACCATGATAGCTCTCCATCAGCTCTTAT 347  
Qy 257 CCAAGACCAATGATGACTACCGGGTGAATGCTTCTGCGGCAACAGGGAATACCCAC 316  
Db 348 CTGAGTGAATAATGATTTACCCCTGGATTTCTACTGACAAATTTTGGACACACCCCA 407  
Qy 317 GCGTGTCTTACCGAATATCC--TGATGACTCTGTGAGCTGTGATCCCTCCATGTGG 373  
Db 408 GCGTGTCTTACCGAATATCC--TGATGACTCTGTGAGCTGTGATCCCTCCATGTGG 467  
Qy 374 ACTGTATCTGAGACCCAGACCTCTTCTTGTCTAAATGAGAAAGGGCCAACTTCATGAG 433  
Db 468 GAAACATATGGGTACCCGACACCTTCTTGTAAAGAAACAGCTTATTTCCACATAG 527  
Qy 434 TGACACGAGCAACAGTACTGGGCTCTTCAAGATGGAAATGCTGTACACATCA 493  
Db 528 CTACTACAGCAAGAAATTCATACGCAATTCATCTTGATCTAATTAAGAGATATA 587  
Qy 494 GCGTACCCCTGATTTTCTGCTGCTGATGACCTCAAGAACTTCCCATGACATCCAGA 553  
Db 588 GACTGACATACCGCTTCTTGTCCGATGATTTGACATATTTCCGATGACCGTCAAT 647  
Qy 554 CGTGACAGATCAGCTTGAGAGCTCATCTGACAGCCCTGCTGCAATCTGTGAC 613  
Db 648 TATGCAATATGAAATCGAAAGTT----- 671  
Qy 614 TTTGAGTTGCTACACATGAAAGACCTGCTGTTGAGTGGCTGGAAGATCTCTGCTG 673  
Db 672 -----TTGGCTTACACATGCGGAGATCCGATACAAATGGAATGAGGGCCCAACTGAG 725  
Qy 674 TCCAAGTGGCTGAGGGGCTGACTGCTGCCCCGATTATCTTGGGATGAGAAGATCTAG 733  
Db 726 TGGGTGTGTGAGGAGTGTCTTGGCGCAATTCAGAGTCTGGGCCACCGGCGAG 785  
Qy 734 GCTGTTTACCAAGCACTACAAACAGGAAATTCACCTGATGAGGTAAGTTTACAC 793  
Db 786 CCATGAGATATTTCTC---TTACGACAGGAAGTACTGCTGCTGCAATGAAATTCAT 842  
Qy 794 TGGAAAGCAATGAGGCTACTATCTGATTCAGATGATGATATCCCAAGCTTACTCATG 853  
Db 843 TTTGAGCTGATGAGGATGATTTAATTAATTCAGATTTAATTCCTGCGCTTAATTTG 902  
Qy 854 TCCGTGCTGCTGCTCTCTGATTCATGATGATGCTGCCCCGCTGCTGAGGCTGAG 913  
Db 903 TTAATCTGCGGATCATTTTGGTTGAATGCAAAATGCGACACTGCAAGGATACACTAG 962  
Qy 914 GCATCACCACCGCTCCTACCATGACACCCAGAGCTGCTGCCGGCTCTTTTGGCTTA 973  
Db 963 GTGTACAAACTGTATTGACATGACAGCTCATGCTGCTCCAGCAATGCGGCTGAGCCA 1022  
Qy 974 AGGTGCTCTAGTGAAGCAATGACATGATGAGTGGCTGTGTCTCTTTGTGTTG 1033  
Db 1023 AGATCTCATATGTAAGTCATGATGATGATGATGATGATGATGATGATGATGATG 1082  
Qy 1034 CTGCTTCTGCTGAGTATGC 1052  
Db 1083 CCAAGTTTACTGAAATATGC 1101

RESULT 12  
US-08-417-330A-19  
: Sequence 19, Application US/08417330A  
: Patent No. 5719057  
: GENERAL INFORMATION:  
: APPLICANT: HADINHAM, KAREN  
: APPLICANT: LE BOURDELLES, BEATRICE  
: APPLICANT: WHITING, PAUL  
: APPLICANT: WINGROVE, PETER  
: TITLE OF INVENTION: STABLY TRANSDUCED CELL LINE EXPRESSING  
: TITLE OF INVENTION: GABA-A RECEPTOR AND NOVEL CLONED  
: NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:  
ADDRESSEE: J. MARK HAND - MERCK & CO., INC.  
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: US  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,330A  
FILING DATE: 05-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: HAND, MARK  
REGISTRATION NUMBER: 36,545  
REFERENCE/DOCKET NUMBER: T11091A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-3905  
TELEFAX: 908-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1866 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding sequence  
LOCATION: 225...1646  
OTHER INFORMATION:  
US-08-417-330A-19

Query Match 9.5%; Score 156.4; DB 1; Length 1866;  
Best Local Similarity 52.0%; Pred. No. 4.3e-36;  
Matches 460; Conservative 0; Mismatches 376; Indels 48; Gaps 3:

Qy 169 GGATATGATGCCAGATTCGGCCCAATTTTAAAGCCCGGAGAGTACCTGCAC 228  
Db 360 GCGTATGACATTCGCTGAGACAGATTTTGGAGGTCCCGCTGCTGGGATGAGAC 419  
Qy 229 ATCTTCAATCAACAGTTTCAGCTCCGTCACAAAGACACATGAGTACCGGGTGAATGTC 288  
Db 420 ATTGACATGCGCAGCATGATATGTTCTGTAAGTCAATATGATATACCTTGACATG 479  
Qy 289 TTTCTGGGCAACAGTGAATGACCCAGCCTGTCTCTACGAGATATCTGATGACTCT 348  
Db 480 TACTTTCAACAAGCTGAGAGATTAAGAGCTGCTCTTAATTAATTAAGT-- 537  
Qy 349 CTGACACCTGATCCCTCATGCTGAGACTCTATCTGGAAGCCAGACACTCTTTGCTAT 408  
Db 538 -TGACTCTGCAACAGAGTGGACAGACGCTGCTGGGCTGATTAATCTTCTGAC 596  
Qy 409 GAGAAAGGGCCCAACTTCCATGAGGTGACACGAGCAACAGTACTGCGCATCTTCAAG 468  
Db 597 GATAAAGATCATTTTGTGACAGAGTGTGCTTAAGAACCGCATGATTCGCTGATCTCT 656  
Qy 469 AATGGGAATGCTGTACAGCATGAGCTGACCCCTCATTTTGTCTGCTGATGAGACTC 528  
Db 657 GATGCAACCGCTCTTATGAGACTCAGAAATCAACACAGCTGCTGATGATGAGACTTA 716  
Qy 529 AAGAACTTCCCATGAGCATGACAGCTGCAACAGTGAAGTGAAGTGAAGTGAAGTGA 588  
Db 717 AGGAGGTACCCAGCTGAGTGAACAAACTGCTTGGAAATGAGAGTA----- 766  
Qy 589 TGCAGCCCTCTGCAATCTCTGCACTTTCAGTTGGCTACACATGAAGACCTGCTGTTT 648





FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 669 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 2..667  
OTHER INFORMATION:  
US-09-002-361-7

Query Match 8.9%; Score 146.6; DB 4; Length 669;  
Best Local Similarity 54.6%; Pred. No. 1.9e-33;  
Matches 367; Conservative 0; Mismatches 260; Indels 45; Gaps 2;

OY 399 CTTTGTATGAGAAAGGGGCACTTCATGAGGTGACACGAGACAGCAACTTACTGCG 458  
DB 1 CTTTGTGAGAAAGCAATCGTACTTCCACAGGCGCACCAAGTAAAGTTCATCGG 60  
OY 459 CATCTTCAAGATGGGAATGCTGTACAGCATCAGGCTGACCTCATTTTGTCTGCGCT 518  
DB 61 CATCCACCACTGGGCTCATCAGCGGTAGCATTAAGCTCACCATCAGCGCTCTCGCCC 120  
OY 519 GATGACCTCAAGAACTTCCCGCATGACATCCAGACGTGACAGATGACAGCTTGAGAGCTC 578  
DB 121 CATGAACCTGCACTTCCCGCATGAGTGGGAGCTGCCACATCGAGATCGAGAGTT- 179  
OY 579 ATCCATACCTCTGACGCCCTCTGCCATCTGTACACTTTGAGTTGGCTACACCATGANAAG 638  
DB 180 -----TCGGCTACACCATGCGGGA 198  
OY 639 CTTGTGTTGAGTGGCTGAGAGATGCTGCTGTCAGAGTGGCTGAGGGGCTGACTCT 698  
DB 199 CATCCGGTCAAAATGAGACGAGGGGNCACACTCGGTGGCGTTTCAACGAAAGTGTGCT 258  
OY 699 GCCCAGTTTATCTTGGGGATGAGAAGATCTAGGCTGTGTACCAAGCACTACAACAC 758  
DB 259 ACCGAGTTCAAGTGTGGGCGCATGTCACGTGGCATGAATATTCGCT--CACAC 315  
OY 759 AGGGAATTCACCTGATGAGTAAAGTTTACCTGGAACGAGAGATGGCTACTATCT 818  
DB 316 AGGAACACTCTCCGCGTGGCGTGGAGATCGATGCGCTGATGGGCTACTACTCT 375  
OY 819 GATTGATGATACATCCCGAGCTCTACTATCGTCAATCCGTGCTGCTGCTGAT 878  
DB 376 GATCCAGATCTACATACCAATCCGGCTCATGCTCAATATATGCTGCTCTTTCTGTT 435  
OY 879 CAACATGATGCTGCCCTGCGCGTGGGCTGGGATCACACCGTGTCAACATGAC 938  
DB 436 GAACCGCAACGAGCGCGCGCGTGGAGTGGGCTGACACCGCTCTCACCATGAC 495  
OY 939 CAACGAGACTGCTGCTCCCGGCTCTTGGCTAAGTGTCTCAATGTAAGGATGCA 998  
DB 496 CACGCTCATGCTTCTCAATATGCGCGCTGCGAAGATCTGTACGTTAAGTCAATGA 555  
OY 999 CATCTGATGCTGCTGCTGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1058  
DB 556 TGTGTACCTGCGACACTGCTGTATGATGTTACACCAAGTGTGAGATACGCGAGGT 615

OY 1059 AAATTTGTTTC 1070  
DB 616 GCGTATATGTC 627

Search completed: June 30, 2003, 23:02:26  
Job time : 116 secs





```
OY 301 CAGTGAATGACCCAGCGCTGTCCTACCGAGAAATATCTGATGACTCTGTGGACCTGCAT 360
DB 301 CAGTGAATGACCCAGCGCTGTCCTACCGAGAAATATCTGATGACTCTGTGGACCTGCAT 360
OY 361 CCCTCCATGCTGGAGCTCTATCTGGAAGCCAGACCTCTTCTTGGCTAATGAGAAAGGGGCC 420
DB 361 CCCTCCATGCTGGAGCTCTATCTGGAAGCCAGACCTCTTCTTGGCTAATGAGAAAGGGGCC 420
OY 421 AACTTCCATGAGAGTGACCCAGCAACAAGTTACTGGCCATCTTAAATAATGGGAATGTG 480
DB 421 AACTTCCATGAGAGTGACCCAGCAACAAGTTACTGGCCATCTTAAATAATGGGAATGTG 480
OY 481 CTGTACACATCAGGCTGACCCCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCC 540
DB 481 CTGTACACATCAGGCTGACCCCTCATTTTGTCTGCTGATGAGACCTCAAGAACTTCCCC 540
OY 541 ATGACATCCAGACGTGACGATGACGCTTGAAGCTCATCTCTGACGCCCTGTG 600
DB 541 ATGACATCCAGACGTGACGATGACGCTTGAAGCTCATCTCTGACGCCCTGTG 600
OY 601 CCATCTCTGTCACTTTCAGTTGGCTACACCATGAAAGACCTGCTGTGAGTGGCTGGA 660
DB 578 -----TTGGCTACACCATGAAAGACCTGCTGTGAGTGGCTGGA 618
OY 661 GATGCTCCTGCTGTCACAGTGAGTGGGGCTGACTGCTGCCAGTTTATCTTGGGGAT 720
DB 619 GATGCTCCTGCTGTCACAGTGAGTGGGGCTGACTGCTGCCAGTTTATCTTGGGGAT 678
OY 721 GAGAAGATCTAAGGCTGTTGTACCAAGCATCAACACAGGGAATTCACCTGCTGACG 780
DB 679 GAGAAGATCTAAGGCTGTTGTACCAAGCATCAACACAGGGAATTCACCTGCTGACG 738
OY 781 GTAAAGTTTACCTTGGAAAGCGGAGATGGGCTACTATCTGATTTACAGTATCCCGAC 840
DB 739 GTAAAGTTTACCTTGGAAAGCGGAGATGGGCTACTATCTGATTTACAGTATCCCGAC 798
OY 841 CTACTCATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
DB 799 CTACTCATGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 858
OY 901 CGTGTGGGCTGCGGATCACCACCGTCTCATCATGACACCCAGAGCTGTGGCTCCGG 960
DB 859 CGTGTGGGCTGCGGATCACCACCGTCTCATCATGACACCCAGAGCTGTGGCTCCGG 918
OY 961 GCCTCTTGGCTTAAAGGTGCTCAGTGAAGCAATCGACATCTGATGGCTGTGCTG 1020
DB 919 GCCTCTTGGCTTAAAGGTGCTCAGTGAAGCAATCGACATCTGATGGCTGTGCTG 978
OY 1021 CTCTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 979 CTCTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
OY 1081 AAGAGATTCATACGACTTTCGAAGAAGGAGGCGCCAAACGCTTGGAGAGATATATC 1140
DB 1039 AAGAGATTCATACGACTTTCGAAGAAGGAGGCGCCAAACGCTTGGAGAGATATATC 1098
OY 1141 CAAGAAGTGTCTTATTTCCGTGCTATGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1099 CAAGAAGTGTCTTATTTCCGTGCTATGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTG 1158
OY 1201 GGAGTCCCAATGGAAGTTTGTGCAATTAATAGTCCCAACCTCCAGCCCTTCTTAAG 1260
DB 1159 GGAGTCCCAATGGAAGTTTGTGCAATTAATAGTCCCAACCTCCAGCCCTTCTTAAG 1218
OY 1261 GAAGAGAAACCCAGCGGAAACTCTAGTGAAGTGAAGCAAGAGAAATTTGACACATCTCC 1320
DB 1219 GAAGAGAAACCCAGCGGAAACTCTAGTGAAGTGAAGCAAGAGAAATTTGACACATCTCC 1278
OY 1321 GGGGCTGTCTTCCCTTCACTTCTCATCTTCAATATCTTCTACTGGTGTCTATATA 1380
DB 1279 GGGGCTGTCTTCCCTTCACTTCTCATCTTCAATATCTTCTACTGGTGTCTATATA 1338
```

```
OY 1381 GTGCTATGCTAGAGATATATCCACCGGCTGTGTGAATAGGATGGAGCTATAGAGCTCT 1440
DB 1339 GTGCTATGCTAGAGATATATCCACCGGCTGTGTGAATAGGATGGAGCTATAGAGCTCT 1398
OY 1441 GCTGTGCTCTGCTGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1399 GCTGTGCTCTGCTGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1458
OY 1501 GTTGGAGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1459 GTTGGAGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1518
OY 1561 GTGGCTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1519 GTGGCTATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1578
OY 1621 TA 1622
DB 1579 TA 1580
```

## RESULT 3

```
US-10-075-846-9
; Sequence 9, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; FILE REFERENCE: D0079 NP
; CURRENT APPLICATION NUMBER: US/10/075,846
; PRIOR APPLICATION NUMBER: 2002-02-13
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 993
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-075-846-9
```

Query Match 49.3%; Score 807.8; DB 9; Length 993;  
Best Local Similarity 90.0%; Pred. No. 1.9e-253;  
Matches 931; Conservative 0; Mismatches 2; Indels 102; Gaps 2;

```
OY 1 ATGACAACTCTTGTCTCTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
DB 1 ATGACAACTCTTGTCTCTGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
OY 61 GTCCCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 61 GTCCCTCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
OY 121 CCCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 CCCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
OY 181 AGGATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 AGGATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
OY 241 AGTTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269
DB 241 AGTTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269
OY 270 -----GGACTACCGGCTGGAATGTCTTCTGCGGCA 300
DB 301 TCTGAAGGCGCATATCTGCAACCTCCAGAGACTACCGGCTGGAATGTCTTCTGCGGCA 360
OY 301 CAGTGAATGACCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 361 CAGTGAATGACCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
```

QY	DB	1178	GCCACTCCTCTGACGAGCAGATGAGAGTCCAAATGGAAGTTCTGGCATTTATAGTCCC	1237
QY	DB	253	GCCACTCCTCTGACGAGCAGATGAGAGTCCAAATGGAAGTTCTGGCATTTATAGTCCC	312
QY	DB	1238	AACCTCCAGCCCTCTTCTAAGGGAAAGAGAAACACGCGGAACCTACGTGACTGAG	1297
QY	DB	313	AACCTCCAGCCCTCTTCTAAGGGAAAGAGAAACACGCGGAACCTACGTGACTGAG	372
QY	DB	1298	CCAAGAGAAATTTGACACCATCTCCCGGGCTCTCTTCCCTTACCTTCCATCTTCATTA	1357
QY	DB	373	CCAAGAGAAATTTGACACCATCTCCCGGGCTCTCTTCCCTTACCTTCCATCTTCATTA	432
QY	DB	1358	TCTTCTACTGAGGTTGCTCTTAATAGTCTATATGTCAGAAATATCCACAGCTCTGTGA	1417
QY	DB	433	TCTTCTACTGAGGTTGCTCTTAATAGTCTATATGTCAGAAATATCCACAGCTCTGTGA	492
QY	DB	1418	TAGGCTGGGAGCTTTAGAGTCTGCTCTGCGCCCTCGATCTCCGCTCGGGTGAGCTTCT	1477
QY	DB	493	TAGGCTGGGAGCTTTAGAGTCTGCTCTGCGCCCTCGATCTCCGCTCGGGTGAGCTTCT	552
QY	DB	1478	CCCTCAGTTTACACTTCATTTAGGGGTTTGGACAGTTCTCTCTGATCTCCACACTAGAACT	1537
QY	DB	553	CCCTCAGTTTACACTTCATTTAGGGGTTTGGACAGTTCTCTCTGATCTCCACACTAGAACT	612
QY	DB	1538	TCAACTACCACTCCCAAAAGCTATGTGGGCTCTATTTGCATATGTGCCAAATGGTGGCTTAC	1597
QY	DB	613	TCAACTACCACTCCCAAAAGCTATGTGGGCTCTATTTGCATATGTGCCAAATGGTGGCTTAC	672
QY	DB	1598	TTATATAA 1604	
QY	DB	673	TTATATAA 679	



Db 177 TCCACCTCAGAACTCACTACACAGTCCCAAGCTATGTGGCTATATGATGATGGCC 236  
Oy 1584 AATGGTGGCTGATCTTAAAGATGGCTATCTACCTA 1622  
Db 237 AATGGTGGCTGATCTTAAAGATGGCTATCTACCTA 275

RESULT 6  
US-10-211-673-11

Sequence 11, Application US/10211673  
Publication No. US20030013158A1  
GENERAL INFORMATION:  
APPLICANT: Le Bourdellies, Beatrice  
TITLE OF INVENTION: HUMAN ALPHA 4 RECEPTOR SUBUNIT  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
City: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/211.673  
FILING DATE: 02-Aug-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB95/02323  
FILING DATE: 29-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Xu, Yang  
REGISTRATION NUMBER: 45,243  
REFERENCE/DOCKET NUMBER: T1292  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 732-594-1307  
TELEFAX: 732-594-4720  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1555 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 47...1402  
OTHER INFORMATION:  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-10-211-673-11

Query Match 11.88; Score 193; DB 9; Length 1555;  
Best Local Similarity 53.38; Pred. No. 4,56-52;  
Matches 486; Conservative 0; Mismatches 380; Indels 45; Gaps 2;

Oy 142 CTACAGCAACTTATGGGCGCAACATCTGATGATGCCAGGATGGCCCAATTTTAA 201  
Db 167 CTCCCACTGAGCGGCGGATGATGCGGTTACCGCCGCACTTCCGGCTGGGATCGGA 226  
Oy 202 GGCCACCCGCGTAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAG 261  
Db 227 GGCCACCCGCGTAACGTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAG 286  
Oy 262 ACCCAATGACTGACCGGCGGATGCTCTTGGGCGCAACAGTGAATGACCCACGCGCTG 321

Db 287 GCCAATAGAGTACACCATGACGGTGTCTCCGACCAAGAGTGGCGGAGACAGGCTC 346  
Oy 322 TCCACGAGAAATATCTGATGATCTCTGACCTGATCTCCATCTGACGACTATAC 381  
Db 347 TCCCTAC--AACCAACCAAGAGACCTTGCGCTGACAGCGCTTGTGGACAACTG 403  
Oy 382 TGAAGCCAGACCTCTTGTGTAATGAGAAAGGCGCAACTTCATGAGGTGACCAAG 441  
Db 404 TGGCTGCCGACACCTTCTATCTGTAAGCCAGAGTGGGCTGTCCAGAGATGACGCTG 463  
Oy 442 GACAACAAGTTACTGCGCATCTTCAGAAATGGAATGCTGTACAGCATCAGGCTGACC 501  
Db 464 GAGAACAAGCTCATCCGCTGACAGCCAGGAGGATGATCTGTACAGCATCCGAATACCC 523  
Oy 502 CTCAATTTGTCCTGCTGATGAGACCTCAAGAACTTCCCATGAGCATCCAGAGTGCAGC 561  
Db 524 TCCACTGTGCGCTGCGATGAGACCTGCGCAAAATTCCTCATGAGCAGAGAGTGCATG 583  
Oy 562 ATGAGCTTGAAGACCTCATCATCTGACAGCCCTGCGCATCTGTCACTTTCAGTT 621  
Db 584 CTGACCTGGAGACT-----AC 601  
Oy 622 GCGTACACATGAAAGACTCTGTGTTGAGTGGCTGGAAGATGCTCTGCTGCCAAGTG 681  
Db 602 GGTACTCATGCGAGACATGCTACTACTGCTGAGAGACCGAGACATCCACGCGG 661  
Oy 682 GCTGAGGGCTGACTCTGCCCCAATTATCTTGGGGATGAGAAGATCTAGCTGTGT 741  
Db 662 CTGACAAAGCTGACGCTGGCGCAGTTCACCATCACAGCTACCGCTTACACAGGAGCTG 721  
Oy 742 ACCAAGCACTACACACAGGAAATTCACCTGCATCGAGGTAAAGTTTCACCTGGAAGCG 801  
Db 722 ATGAACTTCAAGTCCGCTGGCCAGTTCACAGGCTGACGCTGCACCTTCACCTCGGAGG 781  
Oy 802 CAGATGGCTACTATCTGATTCAGATGTACATCCCAAGCTACTCATCATCTCTGCTCC 861  
Db 782 AACCGCGCGTGTATCATCATCATCATCTACATGCTCCCTGCTGCTGCGCATCTCC 841  
Oy 862 TGGCTCTCTTGTGATCAACATGATGATGCTGCCCTCCGCTGGGCGCTGGGATGAC 921  
Db 842 TGGCTCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 901  
Oy 922 ACCGTGCTACACATGAC 981  
Db 902 ACAGTGTGAGATGAC 961  
Oy 982 TACGTGAAGCAATGACATCTGATGATGATGATGATGATGATGATGATGATGATGAT 1041  
Db 962 GCCATCAAGGACACGACGCTACTCTGATGATGATGATGATGATGATGATGATGATGAT 1021  
Oy 1042 CTGAGATATGC 1052  
Db 1022 GTGAGATATGC 1032

RESULT 7  
US-10-239-420-10

Sequence 10, Application US/10239420  
Publication No. US20030096984A1  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION CHANNELS FROM DERMACEPTOR VARIANTS  
FILE REFERENCE: 20629P  
CURRENT APPLICATION NUMBER: US/10/239.420  
CURRENT FILING DATE: 2002-09-23  
PRIOR APPLICATION NUMBER: PCT/US01/09956  
PRIOR FILING DATE: 2001-03-28  
PRIOR APPLICATION NUMBER: 60/193,935  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 10  
LENGTH: 1197  
TYPE: DNA  
ORGANISM: Rhipicephalus sanguineus  
US-10-239-420-10

Query Match 11.4% Score 187.6; DB 9; Length 1197;  
Best Local Similarity 55.0%; Pred. No. 2.3e-50;  
Matches 466; Conservative 0; Mismatches 334; Indels 48; Gaps 3;

```
226 AACATCTCATCAACAGTTTCCAGTCCGCTCCACCAAGACCAACAAAGTACTGGGATAT 285
    ||||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 97 AACATCTTGTAGAGATGCGAGATGATGATGATGATGATGATGATGATGATGATGAT 156
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 286 GTCTTCTTGGCGAACAAGTGAATGACCCAGCCCTGCTTCAACGAAATATCTGATGAC 345
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 157 ATGACGTTCAAGAGCACTGGCGGACAGAGACTCCAGTACGACGATTTGGGGCCGAC 216
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 346 TCTGTGACCTGATCCCTCATGCTGATGATGATGATGATGATGATGATGATGATGAT 405
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 217 GTTCGCTACCTGACGCTCAGCAACCGGACAGAGCTTGGAAAGCCGACCTGTTTCTCC 276
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 406 AATGGAAGGGGCCAATCTTCATGAGTGAACCAAGCAACAAAGTTACTGGGATCTTC 465
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 277 AACGGAAGAGGACACTTCCACAAATCATCATGATCCCAAGTCTCTTACGCAATACAT 336
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 466 AAGATGGGAATGCTGTAGACATCAGGCTGACCTCATATTTTCTGCTGATGAGAC 525
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 337 CCCAAGGGGAGGATCTCTTACGATCAGAAATATCTGGTGTTCATGTCGATGAC 396
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 526 CTCGAACATTCCTCATGACATCCAGACGACGATGACGATGACGATGACGATGACGAT 585
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 397 CTGAATTTTATCTTGTGATTAACAAATCTGCTTATGTCATGATGAGCT----- 448
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 586 CTCGACAGCCCTGCTCAGATCTGTCATCTGATGATGATGATGATGATGATGATGATGAT 645
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 449 -----ATGGGTATACAAAGAGAGACCTGGTG 474
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 646 TTTGATGGCTGGAAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 475 TTTTATG---GAAAGAGGGGATCTGTCAGAGTCAACAAAATCTCAGTGGCAGCT 531
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 706 TTTATCTGGGATGAGAAGATCTAGGCTGTGTACCAACGATACCAACAGAGGAA 765
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 532 TTCA---CGCTGGAAGGTTTCAAAACGACTACTGACCAAGTGAACAACACTGGCAG 588
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 766 TTTACCTGATGAGATGAAAGTTTACCTGGAAGGAGATGGGATGATGATGATGATGAT 825
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 589 TACAGCTGCTGGGCTGAGCTGCTGTTCAAGCGAGATGATGATGATGATGATGATGATGAT 648
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 826 ATGTACATCCAGGCTACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 649 ATCTACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 886 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 945
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 709 ACCCTGATCCCGCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 946 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 769 ATATGAGGATCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 828
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 1006 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 829 ACCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 1066 GTTCTCG 1073
    || | | |
Db 889 GCCTGCG 896
    || | | |
```

RESULT 8  
US-09-969-844-13

Sequence 13, Application US/09969844  
Publication No. US2002019276A1  
GENERAL INFORMATION:  
APPLICANT: Xiao-Zhou Michelle Wang  
APPLICANT: Xavier Georges Sarda  
APPLICANT: Michael David Tomalski  
APPLICANT: Vincent Paul Mary Wingate  
TITLE OF INVENTION: Heliothis glutamate Receptor  
FILE REFERENCE: A32815-1 07267.0178  
CURRENT APPLICATION NUMBER: US/09/969, 844  
CURRENT FILING DATE: 2001-10-03  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 13  
LENGTH: 4621  
TYPE: DNA  
ORGANISM: Heliothis virescens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (764)..  
OTHER INFORMATION: n is a, t, g, or c  
US-09-969-844-13

Query Match 11.4% Score 187; DB 9; Length 4621;  
Best Local Similarity 54.5%; Pred. No. 8.2e-50;  
Matches 472; Conservative 0; Mismatches 346; Indels 48; Gaps 3;

```
208 CCCGTAACGCTGACCTGCAACATCTTCAACAGTTTCAGCTCCGTCACCAAGACCA 267
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 330 CCAGCGGTAGTAGGCTCAATATATTTGTCGAGATATATCAAGATGATGATGATGATGAT 389
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 268 ATGACATCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 390 ATGGAATCTCTGCACTTAACGTTTGGGAAACATGTTAGTGAAGAGGCTCAAAATTC 449
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 328 CGAGATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 450 AATATCTTGGAGTGGCTCAATATCTGACACTGACAGCAAGCAAGATGATGATGATGATGAT 509
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 388 CCAGACCTCTTCTTGTGATGAGAAGGGGCAACTTCCATGAGTGAAGTGAAGTGAAGTGAAG 447
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 510 CCGATCTATCTTCTTCCACAGGAGAGAGGATCTTCCACAAATCATCATGATGATGATGATGAT 569
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 448 AAGTACGCTGATCTTCAAGATGGAATGCTGCTGATGATGATGATGATGATGATGATGATGAT 507
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 570 GTGTACATCCGATTTTCCCAAGGCAAGGCTGATGATGATGATGATGATGATGATGATGATGAT 629
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 508 TTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 630 CTCTGCTGCCCATTAACCTCAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 689
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 568 CTGAGACCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 690 ATGGCTAGT-----ATGGTTGG 707
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 628 ACCATGAAGACCTGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 708 ACCAGAGAGACTTGTATGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 764
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 688 GGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 747
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 765 AACTTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 821
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 748 CACTAACACAGGAGAAATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 822 AAGACTAATACCGGATATACAGTTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 881
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 808 GGTCTATATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 867
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 882 AGTTACTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 941
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Oy 868 TCCTTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```



```
Db 477 CAACGAGAGGAGGATCTTCCACACATCATCATGCCGAGCGATGATCTT 536
Oy 465 CAACAAATGGGAATGCTGTACACATCAGCGCTGACCCCTCTTTGTCCTCCGATGGA 524
Db 537 CCCCAGCGGCAACGCTGACACATCGAATCCCTGAGCGCTCTGCTCCCAAGAA 596
Oy 525 CCGTCAAGAACTTCCCATGACATCCAGACGCTGACAGATGACGCTTGAGAGCTCAT 584
Db 597 CCTCAAGTTATACCCCGGATTAAGACAGACTGCTCGCTAGATAGGCTACTT----- 649
Oy 585 ACTGTCAGCCCTTCCCATCTCTGTCTACTTTGAGTTGGCTACACATGAAAGACTGCT 644
Db 650 -----ATGTTGGACCAACAGAGCACTAGT 674
Oy 645 GTTTAGAGGCTGAGAGATGCTCCTGCTGTCCAAATGAGCTGAGGGGCTGACTGACCCCA 704
Db 675 GTTCTATGAGAGGAGGCGAC---CGGTGAGGTGTGTGAAAACCTTACACCTGCTCG 731
Oy 705 GTTATCTTGGGGATGAGAGATCTAGGCTGTTGTAACAAGCACTAACAGAGGAA 764
Db 732 GT---TCAGCTGGAGAGTTCTCTACAGTACTAGCAAGTAAGACTAATACCGGTGA 788
Oy 765 ATTCACCGCATCGAGTAAGTTTCACTGAGAGCGAGATGGGCTACTATCTGATCCA 824
Db 789 ATACGTTGCTGAGAGTACCTGCTCTTCAAAAGGAGTTCACTTACTGATCA 848
Oy 825 GATGATACATCCCAAGCTCTACTCATGCTGCTGCTGCTGCTGCTGCTGATCAAT 884
Db 849 GATGATACATCCGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908
Oy 885 GAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
Db 909 GGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968
Oy 945 GAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
Db 969 GTCGTCAGGCAATCAACGCGCTCCCTACACAGCGGTGCTCTACACAGAAAGCATTTGATG 1028
Oy 1005 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064
Db 1029 GACTGGTTATGCTCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
Oy 1065 TGTTCTCG 1073
Db 1089 TGCTCTCG 1097

RESULT 11
US-10-239-420-3
; Sequence 3, Application US/10239420
; Publication No. US20030096984A1
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Zheng, Yingcong
; TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION
; TITLE OF INVENTION: CHANNELS FROM DERMACEPTOR VARIANTS
; FILE REFERENCE: 20629P
; CURRENT APPLICATION NUMBER: US/10/239,420
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: PCT/US01/09956
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/193,935
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3442
; TYPE: DNA
; ORGANISM: DERMACEPTOR VARIANTS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (32)...(1225)
US-10-239-420-3
```

```
Query Match 10.5%; Score 171.8; DB 9; Length 3442;
Best Local Similarity 52.9%; Pred. No. 6.5e-45;
Matches 454; Conservative 0; Mismatches 357; Indels 48; Gaps 2;

Oy 202 GGGCCACCGCTGAGACGATGCTGCAACATCTTCAATCAAGCTTTCAGCTCCGATACCAAG 261
Db 197 GGAACGCCACAAAGAGGCTGCAAAATCTACATACGACCTTCCGATCAATCAATCA 256
Oy 262 ACCAATGAGCTACCGGGTGAATGTCTTCTTGGGGACAGAGTGAATGAGCCAGCCGCTG 321
Db 257 GCCAATGAGCTACGAGTGTGATCTTATTTTGGGGACAGCTTGGCAAGATGATGCTTG 316
Oy 322 TCTACCGAGAAATATCCGATGACTCTGTGACCTCGATCCCTCATCTGACTATC 381
Db 317 ACAGGCCCAACGATATCCAGGCCCCGTGACCTCAATGATCAAGCTGTGACAGTATA 376
Oy 382 TGAAGCCAGACCTCTTCTTGTATGAGAAAGGGGCAACTTCCATGAGTGAACCAAG 441
Db 377 TGAAGCCAGAGATATCTTGTGCAAAAGCAAAACAGAGGTTCAATATGATACAGTA 436
Oy 442 GACAACAGTATCTGCGCATCTTCAAGATGGAATGATGCTGTACAGCATCAGGCTGACC 501
Db 437 CCTATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 496
Oy 502 CTGATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
Db 497 CTGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 556
Oy 562 ATGAGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
Db 557 ATGAGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
Oy 622 GGTACACATGAAAGACCTGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 681
Db 581 -----ACAACGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 631
Oy 682 GCTGAGGGCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
Db 632 TTTGAGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 688
Oy 742 ACCAAGCTACACACAGGGAATTCACCTGATGATGATGATGATGATGATGATGATGATGAT 801
Db 689 ATGAGAAATTTTACATGCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 748
Oy 802 CAGATGGCTACTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 861
Db 749 TCACTGGGCTACACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 808
Oy 862 TGGCTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 921
Db 809 TGGCTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868
Oy 922 ACCGCTCTACCATACCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 981
Db 869 ACCGCTCTACCATATTTTCCAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 928
Oy 982 TACGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1041
Db 929 TACGTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 988
Oy 1042 CTGAGATATGCTGCCATAA 1060
Db 989 CTGAGATTACCGCTGCTCA 1007

RESULT 12
US-09-815-925-10
; Sequence 10, Application US/09815925
; Patent No. US20020127199A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
```

```
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Yang, Yonghong
APPLICANT: Zhang, Jie
APPLICANT: Wehrman, Tom
APPLICANT: Drmanac, Radote T.
TITLE OF INVENTION: No. US20020127199A1el Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/815,925
CURRENT FILING DATE: 2001-03-22
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PL_FL_genes Version 2.0
SEQ ID NO 10
LENGTH: 1467
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (46)..(1464)
US-09-815-925-10
```

Query Match 10.3%; Score 168.6; DB 10; Length 1467;  
Best Local Similarity 52.6%; Pred. No. 4.3e-44;

Matches 474; Conservative 0; Mismatches 379; Indels 48; Gaps 3;

```
OY 169 GGATATGATGCCAGATTGGCCCAATTTAAAGCCCAACCGGTGACGTGACCTGAC 228
DB 184 GGTACAGACATTCGCGTAAAGCCGACTTCGGGGGTGCCCCGGCTGCGGGGATGAAC 243
OY 229 ATCTTCAACAGTTTCAGTCCGACCAAGACACCAATGACATGACGGGTGATGTC 288
DB 244 ATGACATGCGCCGACATGACATGTTTCCGAGTCAACATGATTAACCTTAACCATG 303
OY 289 TTCTTGGGCAACAGTGAATGACCCAGCCTGTCTTACCGAATATCTGATGACTCT 348
DB 304 TATTTCAACAAATTTGAGAGATTAAGGCTGCTATTTGATGCTTCCCTCAGC--- 360
OY 349 CTGGACCTGATCCCTTCATGCTGACCTATCTGGAAGCCAGACCTCTTTTGTAT 408
DB 361 CTCACCTTGACATGAGTGGCTGACACAGCTATGGTCCCGACACATATTTCTTAAT 420
OY 409 GAGAAAGGGGCCAATTCATGATGAGTACCGACCGAACAAGTACTGGGCATCTCAAG 468
DB 421 GACAAAAGTCATTTGTGATGATGAGTGAAGTAAACCGCATGATCCGCTTTCACCT 480
OY 469 AATGGAAATGCTGTACAGCATCAGGCTGACCCCTATTCTGCTGCTGATGACCTC 528
DB 481 GATGGACAGTGTGTATGAGGCTCAGATACACACAGACAGCATGATGATGAGCTC 540
OY 529 AAGACTTCCCATGACATCCAGACGTGACAGTACCTGAGAGCTGATCCATACCTC 588
DB 541 AGAGAGTACCCCTCGAGCAGCAAGACTCCTGGAATTTGAAAGCT----- 589
OY 589 TGCAGCCTCTGTCATCTGTGACATTTGATGATGATGATGATGATGATGATGATG 648
DB 590 -----ATGGCTACACAGGATGATGATGATGATGATGATGATGATGATGATG 618
OY 649 GAGTGGCTGGAAGATCTCTGCTGTCAAGTGGCTGAGGGGCTGACTGCTGCTGCTG 708
DB 619 TACTGCGAGGGGAGCAAGGCTGTACCGAGTGAAGAGATTGAGTCCCGCAGT-- 676
OY 709 ATCTTCCGGATGAGAGATCTAGCTGTGTACCAAGCATACACAGGAAATTC 768
DB 677 -TCTCATGCTGAGAGACCGTCTGCTGAGAGAAATGTTCTTCCGACAGGTGCTAT 735
OY 769 ACCTGATGAGGTAAGTTTCACTGGAACCGCAGATGAGGCTACTATGATTCAGAG 828
DB 736 CCTGACTGTGACTGACTTTCGTTGGAAGAGAACTGATGATGATGATGATGATGAT 795
```

```
OY 829 TATATCCAGGCTTACTCATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
DB 796 TATATCCCTTATATGATTAAGATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 855
OY 889 GCTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
DB 856 GATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 915
OY 949 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008
DB 916 ACCCAGCTTTCGAGAGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 975
OY 1009 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
DB 976 ATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1035
OY 1069 T 1069
DB 1036 T 1036
```

## RESULT 13

US-10-239-420-1

Sequence 1, Application US/10239420

Publication No. US20030096984A1

GENERAL INFORMATION:

APPLICANT: Cully, Doris F.

APPLICANT: Zhang, Yingsong

TITLE OF INVENTION: DNA MOLECULES ENCODING LIGAND GATED ION

FILE REFERENCE: CHANNELS FROM DERMACENTOR VARIABILIS

CURRENT APPLICATION NUMBER: US/10/239,420

PRIOR FILING DATE: 2002-09-23

PRIOR APPLICATION NUMBER: PCT/US01/09956

PRIOR FILING DATE: 2001-03-28

PRIOR APPLICATION NUMBER: 60/193,935

PRIOR FILING DATE: 2000-03-31

NUMBER OF SEQ ID NOS: 13

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 3598

TYPE: DNA

ORGANISM: Dermalcentor variabilis

FEATURE:

NAME/KEY: CDS

LOCATION: (170)..(1363)

US-10-239-420-1

Query Match 10.2%; Score 167; DB 9; Length 3598;

Best Local Similarity 52.5%; Pred. No. 2.5e-43;

Matches 451; Conservative 0; Mismatches 360; Indels 48; Gaps 2;

```
OY 202 GGGCCACCGGTGACCTGACATCTGCAACATCTTCAACAGTTTCAAGTCCGTCAAC 261
DB 335 GGAAAGCCCAACAAAGTGGCTTGGAAATCTACATGAGAGCTTCCGGTCAATAATCCA 394
OY 262 ACCCAATGAGTACCGGGGATGATGCTTCTTGGCCAGACAGTGAAGTACCCAGCCG 321
DB 395 GCCACATGACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 454
OY 322 TCTTACCGAGAAATATCTGATGATGATGATGATGATGATGATGATGATGATGAT 381
DB 455 ACAGCCCAAGTATTCAGAGCCCTGGAACCTCAATGATTCGAAGCTGGGAGCGTATA 514
OY 382 TGAACCCAGACTCTTTCTTTGCTATGAGAAAGGGGCCAATCTTCATGATGATGAT 441
DB 515 TGAAGCCGGAATATCTTCTTCAATGCAAAACAGCAGAGTTCGAATATGTCACAGTA 574
OY 442 GACAAACAGTACTGCGATCTTCAAGAAATGGAATGTGTGATGATGATGATGATGAT 501
DB 575 CCTAATGTACTGTCCTGCTTAAACCGGAAGAAATTTCTATATGATGATGATGATGAT 634
```





**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2003, 20:56:06 ; Search time 2191 Seconds  
(Without alignments)  
12122.592 Million cell updates/sec

Title: US-10-075-846-3  
Perfect score: 1640  
Sequence: 1 atgacacactctgtctctgc.....taaaaaaaaaaaaaaaaaa 1640

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
EST:  
1: em\_estlba:\*  
2: em\_estlum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	401.8	24.5	662	9	AU169868 AU169868
2	302.4	18.4	817	12	BG404477 BG404477
3	295.6	18.0	678	10	AV729257 AV729257
4	292.4	17.8	779	14	B0443713 B0443713
5	266.6	16.3	692	12	BE981841 BE981841
6	255.2	15.6	720	12	BE981459 BE981459

Result	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	ALIGNMENTS
1	AU169868	AU169868	AU169868	AU169868	EST	Japanese medaka.	1 (bases 1 to 662)	Mita, K., Ishikawa, Y. and Yamauchi, M.	Establishment of cDNA database of medaka, <i>Oryzias latipes</i>	Unpublished (2001)	Contact: Mita K Genome Research Group National Institute of Radiological Sciences Anagawa 4-9-1, Inage, Chiba 263-8555, Japan Email: kmila@nirs.go.jp method: uni-directional sequence direction: sequenced from T3 primer (5' -> 3')	Location/Qualifiers 1..662 /organism="Oryzias latipes" /strain="HNI"	BM440664 pcr.in.pk0 BC027094 Mus muscu BC022502 Homo sapi BF906462 IL0-OT012 B1736322 603360026 AL538200 AL538200 B1825190 603072067 AM280976 f148d02.y B0938794 AGENCOURT AK013727 Mus muscu BC189196 RST8236 A BE830621 PM2-MT002 AU133223 AU133223 BC029850 Homo sapi BB653397 BB653397 B0044592 UT-M-EO- B0042344 UT-M-EO- AL229643 Tetradon B0067799 AGENCOURT BG186166 RST5011 A BG203115 RST22484 BG193287 RST12516 BE830616 PM2-MT002 BG196871 RST16098 B1913857 603183256 F06915 HSC10G061 n BM129116 f177906.y B1736340 603360049 B186165 RST5010 A AL247390 Tetradon B1552703 603193953 BM964113 UT-M-EO- BM587895 170006873 B1535004 603193642 BG404434 602420552 F13483 HSC3D6111 n B1392407 p9p1n.pk0 R13309 yf75d07.r1

```

/db_xref="taxon:8090"
/clone="br5332"
/clone_lib="OI-br-ad CDNA"
/sex="female/male mixed"
/tissue_type="brain"
/dev_stage="adult"
BASE COUNT      155 a      205 c      164 g      138 t
ORIGIN

Query Match      24.5% Score 401.8; DB 9; Length 662;
Best Local Similarity 76.4%; Pred. No. 2e-74;
Matches 540; Conservative 0; Mismatches 122; Indels 45; Gaps 2;

OY 195 TTTTAAAGGCCCCCGCTGAACGTGACCTGCAACATCTTTCATCAACAGTTTCAGTCCGT 254
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 TTTTAAAGGTCCACCTGTAAAGCTCAACCTGCAACATTTTATCAACAGTTTGGATTCAT 60

OY 255 CACCAAGACCCACATAGACTACCGGGGTGATGTTCTTTGCGGCAACAGTGAATGACCC 314
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CGCTGAACACACCATGAGCTACAGAGTACATCTTCCGAGCAGCATGTGAGACGCC 120

OY 315 AGCGTGTCTACGAGATATCTGATGACTCTGACCTCGACCTCGATGCCCTCATGCTGGA 374
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 121 CGGCTGGCTACAGCGATATCCGAGACTGCTGATCTGAGACCCCTCATGTTGGA 180

OY 375 CTCTATCTGGAAGCCAGACCTCTCTTCTGCTAATGAGAAAGGGGCCAATTCATGAGT 434
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 181 CTCACTCTGGAAGCCGAGATCTGTTCTTCTAATGAGAAAGGGGCCAATTCACAGAGT 240

OY 435 GACCAAGCAACAAGTACTGCGCATCTTCAAGAAATGGGAATGTGCTTACAGATCG 494
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 241 CACCAAGCAACAAGTACTGCGCATCTTCAAGAAATGGGAATGTGCTTACAGATCG 300

OY 495 GCTGACCCATCTTGTCTGCTGCTGATGAGCTCAAGAACTTCCCATGAGATCCAGAC 554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 AATCACTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

OY 555 GTGACGATGACAGCTTGAAGCTCATCTCTGACAGCTCTGCTGCTGCTGCTGCTGCTG 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 CTGACATCAATGACAGCTGAGAGCT----- 383

OY 615 TTCAGTTGGCTACACCATGAAAGACCTGCTGTTGTTGAGTGGATGATGCTCTCTGCT 674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 384 -----TCGGCTACACCATGAAAGACCTGCTGTTGTTGAGTGGATGATGCTCTCTGCT 435

OY 675 CCAAGTGGCTGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 436 GAGAGTGGCTGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495

OY 735 CTGTTTACCAAGCACTACAAACAGAGGAATTCACCTGCTGATGAGTAAAGTTTACCT 794
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 496 CTGACTCCACCAAGCACTACAAACAGAGGAATTCACCTGCTGATGAGTAAAGTTTACCT 555

OY 795 GGAAGGCAAGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 854
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 556 GGAAGGCAAGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615

OY 855 CCGTCTCTGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 616 CCGTCTCTGAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 662

RESULT 2
BG404477      817 bp      mRNA      linear      EST 12-MAR-2001
LOCUS      602420604F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4527768 5',
DEFINITION      mRNA sequence.
ACCESSION      BG404477
VERSION      BG404477.1 GI:13297925
KEYWORDS      EST.
SOURCE      house mouse.
ORGANISM      Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS      1 (bases 1 to 817)
TITLE      NIH-MGC http://imgc.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM10437 row: 9 column: 01
High quality sequence start: 6
High quality sequence stop: 808.
Location/Qualifiers
1. 817
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4527768"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site:1; Note:
Site-2: Sail; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT      223 a      190 c      192 g      212 t
ORIGIN

Query Match      18.4% Score 302.4; DB 12; Length 817;
Best Local Similarity 71.6%; Pred. 1.5e-53;
Matches 496; Conservative 0; Mismatches 180; Indels 17; Gaps 7;

OY 697 CTGCCCCATTTATCTGCGGATGAGAGATCTAGGCTGTTGACCAAGACTACAAAC 756
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 113 CTGCCCCATTTATTTTGAAGAGAGAGAGAGCTGTTATTTGCAACAAAGACTTACAAAC 172

OY 757 ACAGGAAATTCACCTGCATGATGAAAGTTTCACTGAGGAGGAGGAGGAGTACTAT 816
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 173 ACTGGCAATTTACCGCATGAGTGGTCAAGTTTCACTGAGGAGGAGGAGTACTAT 232

OY 817 CTGATTCAGATGATATCCCAAGCTTCACTGATGATGATGATGATGATGATGATGATG 876
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 233 TTGATTCAGATGATATCCCAAGCTTCACTGATGATGATGATGATGATGATGATGATG 291

OY 877 ATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 936
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 292 ATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 350

OY 937 ACCACCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 996
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 351 ACTAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 409

OY 997 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1056
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 410 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 469

OY 1057 ATTAATTTGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 470 GTGAA-CTTGTTCTCCAGGCACTAAGAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTG 528

OY 1117 CAAGCTTGGAGAGATATCAATCAAGAAAGTGTCTTATTTCCGTTGCTATGCTTGG 1176
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 529 CAGATTAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 588

OY 1177 GGCACATGCTGAGGAGCAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1236
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 589 GGTCACTGCTCCAAATGAAAGATG-----GCACAGCTGTCAAGGCTACACCTGC 638

```

QY	1237	CAACCTCGAGGCCCCCTCTTCTTAAGGAGAGAAACCAAGCGGAACACTCTAGGTGACTGA	1296
Db	639	CAACCCACTTCTACACAAACCCCAAGAGATGCAGATCTTTCAGAGAAAGAT-TGTGGATGG	697
QY	1297	GCCAGAAGAAATTGACACCATCTCCCGGGCTGTCTCCCTTCACCTTCCCTCATCTCAAT	1356
Db	698	GCAAAAAGAAATTGACACCATATCTCGAGACTCTTCCCACT--GGCTTCCTCATTTTCAAC	755
QY	1357	ATCTTCTACTGGGTGTCTATPAAAGTGGTAAAGG	1389
Db	756	ATCTTTACTGTGATCACATACAAATCATTTGGG	788

RESULT 3	AV729257	LOCUS	AV729257	678 bp	mRNA	linear	EST 17-OCT-2000
DEFINITION	AV729257	HTC	Homo sapiens	CDNA	clone	HTCHB03.5',	mRNA sequence.
ACCESSION	AV729257						
VERSION	AV729257.1	GI:10838678					
KEYWORDS	EST.						
SOURCE	human.						
ORGANISM	Homo sapiens						

REFERENCE	1 (bases 1 to 678)
AUTHORS	Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
TITLE	Homo sapiens cDNA HTC clones
JOURNAL	Unpublished (2000)
COMMENT	Contact: Zeguang Han

TITLE	JOURNAL	COMMENT
Homo sapiens CDNA HTC clones unpublished (2000)		Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex. 45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.

FEATURES	SOURCE	Location/Qualifiers
	1. .678	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="HRCHE03"	
	/clone_lib="HRC"	
	/tissue_type="Hypothalamus"	
	/dev_stage="Adult"	
	/lab_host="SOLR"	
	/note="Vector: plnuascript sk(-); Site_1: ECORI; Site_2: XhoI."	
BASE COUNT	199 a 148 c 150 g 160 t	
ORIGIN		1 others

	Query Match	18.0%	Score 295.6	DB 10	Length 678
	Best Local Similarity	74.9%	Pred. No. 4.3e-52		
	Matches	370	Conservative	0	Mismatches 124; Indels 0; Gaps 0;
OY	84	AAAGAGGAGCTCAAACTGTGGAACCAAGGGGTCCCGAGCCATGTCCTCCCTGTGATTTCT	143		
Db	49	AGACCATGACATCCAGGCTGTGGAAAAACAACCTTCACAGACCCTATCTCCTCAGATTTCTT	108		
OY	144	AGACAAATTTATGCGGCGCAACATCGATGATGATGATGCGGAGTATGGGCCCATTTTAAAGG	203		
Db	109	GGACAACTTAATGGGAAAGGACATCAGGATATGATGCAAGAAATCAAGCCAAATTTTAAAGG	168		
OY	204	CCCAACCCGTGAACGTGACCTTCAACATCTTCATCAACAGTTTCAGCTCCGTACCAACAG	263		
Db	169	TCTCTCAATAAACGTTACTTCATATATTTTATCAACAGTTTGGATCATCAGAAAC	228		
OY	264	CACATGTGACATACCGGATGATGCTCTTTCGCGCAACAGTGAAATGACCCACGCTGTC	323		
Db	229	GACCATGACATACCGAGTAAATTTTTTCTAGACAACAGTGAAATGATTCACGGCTGTC	288		

QY	324	CTACCGGAATATCTCTATGACTCTGTGGACCTCGATCCCTCAATGGTGACCTATCTG	383
Db	289	GTCAGTGTGATGACCATGATGACTCCCTGTGGACTTGTGGACCATCATGCTAGACTCATTTG	348
QY	384	GAAGCCAGACCTCTCTTTTGTCTAATGAGAAAGGGCCCACTTCATGAGGTGACCAAGCA	443
Db	349	GAAGCCAGATTTCTTTTGGCAATGAGAAAGGTGCCCAACTTCCAGCATGTCCACACATGA	408
QY	444	CAACAGTGTACTGGGCATCTTCAAGATGGAATGTGCTGTACAGCATCAGCGTGACCCCT	503
Db	409	CACACAAATTTGCTACGCGATTTGGAAAAATGGCAAAAGTGTCTTACAGTATCAGACTTCACCTT	468
QY	504	CATTTTGTCTGCTGTATGGACCTTCAAGAACTTCCCATGAGACATTCAGAGCTGACAGAT	563
Db	469	GACCTTATCCGTGCTCCCATGACTTGAAGAACTTCCGATGGATGTCCAGACCTGTACAAT	528
QY	564	GCAGCTGGAGAGCT	577
Db	529	GCAGCTGGAGAGTT	542

[illegible]

REFERENCE 14 (bases 1 to 779)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Dr. James Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: MCC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 This clone was contributed by the Brain Molecular Anatomy Project  
 (BMAP)  
 Seq primer: pXX-5.

```

FEATURES
source
Location/Qualifiers
1. 779
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone_image="5708257"
/clone_lib="NIH_BMAP_EMO"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pIX-Asc; Site_1: Ecor I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with Not I, and then cloned directionally into pIX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTCGCGGCAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institutes of Mental Health

```

BASE COUNT 222 a 167 c 165 g 225 t (NIMH), Hemin Chin, Ph.D., program coordinator."

Query Match 17.88; Score 292.4; DB 14; Length 779;  
Best Local Similarity 74.58; Pred. No. 2e-51;  
Matches 368; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

84 AAAAGAGAGTCAATCTGGAACCAAGGGGTCGCCCATGTCCTCTGATTCCT 143  
197 AGACCATGACTCCAGCTGTGGAAAACATCCCTGCAACCCCTATCTCTTACGATTTCTT 256  
144 AGACAACTTATGAGGGGCAACATCTGATATGATGCGAGATTCGCCCAATTTTAAAG 203  
257 GGATTAATTAATGAGGAGAACATCAGATGATGATGCAAGATCAGGCCAATTTTAAAG 316  
204 CCCACCCGTAAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCAAGAC 263  
317 TCCCTCAGTAAACCTTACTTGCATATTTTATCAACAGTTTGATCAGTACAGCAAGAAC 376  
264 CACATGAGCTACCGGGGATGATGCTCTTGGGCAACAGTGAATGACCCAGCCTGTC 323  
377 CACATGAGCTACCGAGTGAACATTTTCTGAGACAGCAGTGAATGATTCAGCGCTGCG 436  
324 CTACCGGAATATCTGATGATGATCTCTGACCTGATCCCTCATGCTGATCTTATCTG 383  
437 ATACAGTACGATACCAATGATTCCTCGATTTGGATCCCTCATGTTGATTCGATTTG 496  
384 GAAGCCAGACCTCTTCTTCTGATGAGAAAGGGGCAACCTTCATGATGATGACCAAGCA 443  
497 GAACCGGATTTGTTCTTCCCAATGAAGGGGCAATTTCCATGATGATGACCAAGCA 556  
444 CAACAACTGATGCGGATCTTCAAGATGGAATGCTGCTACAGCATGACGCTGACCT 503  
557 CACAAGTTGTTGGGATTTCCAAAATGCAAAAGTCTCTACAGTATTGACGACCTT 616  
504 CATTTTCTCTGCTGATGAGCTCAAGAACTTCCCATGAGATCCAGACGTCGACGAT 563  
617 GACTTATCTTCTGCAATGACCTGAGAACTTCCCAATGATGATCCAGACGCTGACAT 676  
564 GCAGCTGAGAGCT 577  
677 GCAGCTGAGAGCT 690

RESULT 5  
BE981841 692 bp mRNA linear EST 29-APR-2002  
LOCUS BE981841  
DEFINITION UT-M-CGDP-bdd-h-07-0-UI.s1 NIH\_BMAP\_Ret4\_S2 Mus musculus cDNA clone  
UT-M-CGDP-bdd-h-07-0-UI 3', mRNA sequence.  
ACCESSION BE981841  
VERSION BE981841.1 GI:10651376  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 692)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: meste@mail.nih.gov  
Oligo-dT track not found. Not 1 site shown in beginning of sequence  
is likely internal to the message. cDNA library preparation: M.B.  
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA

clones from RESEARCH GENETICS. It should be noted that Bento Soares  
is generating a small number of additional specialized  
non-redundant arrays of BMAP cDNAs whose availability will be  
considered under appropriate and limited collaborative arrangements  
The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
Yale University School of Medicine  
Seq primer: M13 Forward  
POLYA-No.

FEATURES  
source location/Qualifiers  
1..692  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UT-M-CGDP-bdd-h-07-0-UI"  
/clone\_1lb="NIH\_BMAP\_Ret4\_S2"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not I; Site 2: Eco RI; The  
NIH\_BMAP\_Ret4\_S2 library is a subtracted library,  
ultimately derived from mouse retina tissue libraries at  
various stages of development. For a detailed description  
of the library from which this clone was derived, please  
visit our web site at brainest.eng.yuova.edu. The tissue  
for this library was contributed by Dr. Xin-Yuan Fu, Yale  
University School of Medicine  
TAG\_SEQ=None found"

BASE COUNT 190 a 159 c 144 g 198 t 1 others  
ORIGIN

Query Match 16.38; Score 266.6; DB 12; Length 692;  
Best Local Similarity 70.98; Pred. No. 5.4e-46;  
Matches 353; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

80 TGGCAAAAGGAAAGTAAATCTGGAACCAAGGGGTCGCCCATGTCCTCTGATTCCT 139  
172 TGGCCACAAAGGAAACAAACAGTGGGATCTGCAATGCTCAATGCTGATTCGATTC 231  
140 TCTAGACAAACTTATGGGGGCAACATCTGATATGATGCGAGATTCGCGCAATTTTA 199  
232 TTCTGACAAACTAATGGGAGGAGCATGGGGATATGATGCAAGATCAGACCAACTTCA 291  
200 AAGGCCCAAGGAGAGCTGACCTGCAACATCTTCAATCAACAGTTTACGCTCGTCA 259  
292 AAGTCTCTCCAGTAAATGTCACATGCAACATATTCATAAAGCTTGGCTCCATTTG 351  
260 AGACCAACATGAGTACCGGGTGAATGCTCTTGGCGGCAACAGTGAATGACCAAGCC 319  
352 AGACGACTATGATTTACAGAGTAACTTTTCTGCTCAGAAAGTGAATGATTCCTGTC 411  
320 TGTCTACCGAGATATCTGATGACTCTGACCTGATCCCTCATGCTGAGACTCTA 379  
412 TTGCATGACGATGATATCTGACGATTCATAGACCTGACCGGTATGTTGATTCGA 471  
380 TCTGGAAGCCAGACCTTTCTTGTCTATATAGAAAGGGGCAACCTTCATGAGTGACCA 439  
472 TTTGGAACCTGACTGTTCTTGTCTATATAGAAAGGGGCAACCTTCATGAGTGACCA 531  
440 CGGACCAAGTACTGCGCATCTTCAAGATGGAATGCTGTACAGCATGAGGCTGA 499  
532 CAGATTAACAGCTGCTAGAAATTTTCANAAATGGAATGCTTTTATCAATAGGTTGA 551  
500 CCTCATTTTCTCTGCTGCTGACCTCAAGAACTTCCCATGAGCATCAGACGCTGA 559  
592 CATTAAACACTCTCTGCTCAATGATTCGAAGATTTCCCAATGAGATGTCAGAAACATGCA 651  
560 CGATGACGCTTGAAGCT 577  
652 TAATGCAACTTGAAGCT 669

RESULT 6  
BE981459 720 bp mRNA linear EST 29-APR-2002  
LOCUS BE981459

DEFINITION UI-M-CG0P-bdc-d-08-0-UI.s1 NIH\_BMAP\_Ret4\_S2 Mus musculus cDNA clone  
 UI-M-CG0P-bdc-d-08-0-UI 3', mRNA sequence.  
 ACCESSION BE981459  
 KEYWORDS BE981459.1 GI:10650587  
 SOURCE EST.  
 ORGANISM house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclerognathii; Muridae; Murinae; Mus.  
 1 (bases 1 to 720)  
 REFERENCE Ronald,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mesf@mail.nih.gov  
 Oligo-dT track not found, Not 1 site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
 clones from RESEARCH GENETICS. It should be noted that Bento Soares  
 is generating a small number of additional specialized  
 non-redundant arrays of BMAP cDNAs whose availability will be  
 considered under appropriate and limited collaborative arrangements  
 The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
 Yale University School of Medicine  
 Seq primer: M13 Forward  
 POLYA-No.

FEATURES  
 source  
 location/Qualifiers  
 1..720  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-CG0P-bdc-d-08-0-UI"  
 /clone\_lib="NIH\_BMAP\_Ret4\_S2"  
 /lab\_host="DH10B (Life Technologies)"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
 NIH\_BMAP\_Ret4\_S2 library is a subtracted library,  
 ultimately derived from mouse retina tissue libraries at  
 various stages of development. For a detailed description  
 of the library from which this clone was derived, please  
 visit our web site at brainest.eng.uiowa.edu. The tissue  
 for this library was contributed by Dr. Xin-Yuan Fu, Yale  
 University School of Medicine  
 TAG\_SEQ=None found"

BASE COUNT 198 a 162 c 151 g 206 t 3 others  
 ORIGIN

Query Match 15.6%; Score 255.2; DB 12; Length 720;  
 Best Local Similarity 70.7%; Pred. No. 1.3e-43;  
 Matches 352; Conservative 0; Mismatches 145; Indels 1; Gaps 1;

QY 80 TGGCAAGAGAGATCAATCTGGAACCAAGGGTCCAGCCCATGTCCTTGATT 139  
 171 TGGCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 230  
 140 TCTGTAGACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 199  
 231 TTCTGTAGACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 290  
 QY 200 AAGCCCAACCCGTGAGCTGACCTGCAACATCTTCATCAACAGTTTCAGCTCCGTCACCA 259  
 291 AAGGTCTCTCCAGTTATGTACATGCAACATATTCATAAAGCTTTGGCTCCATTCAG 350  
 DB 260 AGACCAATGAGTACAGCGGGTGAATGCTTCTTGGCGCAACAGGGAATGACCCAGCC 319  
 351 AGACCAATGAGTACAGCGGGTGAATGCTTCTTGGCGCAACAGGGAATGACCTCGTC 410

QY 320 TGTCTACCGAGATATTCGTGATGCTCTGACCTCGATCCCTTCAGTGCATCTGA 379  
 411 TTGCATACAGTAATATCTCTGACGATTTATTAGACCTTGACCCGCTATGTTGATTCCA 470  
 QY 380 TCTGGAAGCCAGACCTCTCTTCTTCTTAATGAGAAAGGGGCCAATCTTCATGAGTGACCA 439  
 471 TTTGGAAACCTGACCTGTTGTTCTTCTTAATGAGAAAGGGGGCTTAATCCAGCAAGTCACCA 530  
 QY 440 CGGAAACAAGTACTGCGCATCTTCAGAAATGCTGCTACAGCATCAGGCTGCA 499  
 531 CAGATTAACAAGCTGCTTAAGAAATTTTCANAAAGAAATGCTTTTAAATTAAGGTTGA 590  
 QY 500 CCCATATTTTGTCTGCGCTGATGACCTCAAGAACTTCCCATTCGACATCCAGAGCTGCA 559  
 591 CATTAACACTCTCCCTGCTCCATGATGATCTCAGAA-TTTCCAATGATGATACACATGCA 649  
 QY 560 CGATCGACCTTGAGAGCT 577  
 650 TAATGCAACTTGAAAGCT 667  
 DB

RESULT 7  
 BM440664  
 LOCUS 552 bp mRNA linear EST 01-FEB-2002  
 DEFINITION (pgrln.pk003.m19 Normalized Chicken Reproductive Tract cDNA Library  
 (pgrln) Gallus gallus cDNA clone pgrln.pk003.m19 5' similar to  
 p114A99701A9970 glycine receptor alpha-4 chain - mouse (fragment  
 ), mRNA sequence.  
 BM440664  
 ACCESSION BM440664.1 GI:18471439  
 VERSION EST.  
 KEYWORDS chicken.  
 SOURCE ORGANISM  
 Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 552)  
 COGburn,L.A. and Nys,Y.  
 ESTs from Normalized Chicken Reproductive Tract cDNA Library-  
 University of Delaware and INRA, Tours-Poultry Unit Project  
 Unpublished (2002)  
 JOURNAL Contact: Larry A. Cogburn  
 University of Delaware  
 Townsend Hall, Newark, DE 19717, USA  
 Tel: 302-831-1335  
 Fax: 302-831-2822  
 Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES  
 source  
 location/Qualifiers  
 1..552  
 /organism="Gallus gallus"  
 /strain="Commercial broiler and layer"  
 /db\_xref="taxon:9031"  
 /clone="pgrln.pk003.m19"  
 /clone\_lib="Normalized Chicken Reproductive Tract cDNA  
 library (pgrln)"  
 /sex="Male and Female"  
 /tissue\_type="Testis, ovary and oviduct"  
 /dev\_stage="Various stages: embryonic, post-hatch, immature  
 and sexually-mature"  
 /lab\_host="E. coli EMDH10B"  
 /note="Vector: PCMVSPORT6; Library made from three total  
 RNA pools from each tissue (testis 25%, ovary 25%,  
 oviduct 50% of final RNA pool); Single pass sequencing  
 from 5'-end"

BASE COUNT 85 a 198 c 181 g 88 t  
 ORIGIN

Query Match 14.6%; Score 239.6; DB 13; Length 552;  
 Best Local Similarity 76.9%; Pred. No. 2.7e-40;  
 Matches 306; Conservative 0; Mismatches 89; Indels 3; Gaps 1;

QY 19 GCACCCCT 78





Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL06t2-IL0-OT0123-081200-500-809&t3=2000-12-08&t4=1)

Seq primer: puc 18 forward  
High quality sequence stop: 518.

## FEATURES

source

Location/Qualifiers

1. 536

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="OT0123"

/dev\_stage="Adult"

/note="Organ: ovary; Vector: puc18; Site.1: Sma1; Site.2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent Application No. 196

716 - Ludwig Institute for Cancer Research) profiles

into the puc 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

stringency conditions."

## BASE COUNT

118 a 144 c 133 g 141 t

118 a 144 c 133 g 141 t

118 a 144 c 133 g 141 t

118 a 144 c 133 g 141 t

118 a 144 c 133 g 141 t

118 a 144 c 133 g 141 t

## Query Match

13.1%; Score 215.6; DB 12; Length 536;

## Best Local Similarity

67.7%; Pred. No. 3e-35;

## Matches

321; Conservative 0; Mismatches 144; Indels 9; Gaps 1;

## ORIGIN

857 TGTCTGGGTCCTCTTGTGATCAATGATGCTGCCCCCTGGGCTGGGCA 916

## ORIGIN

532 TCTCTGGGTCCTCTTGTGATCAATGATGCTGCCCCCTGGGCTGGGCA 473

## ORIGIN

917 TCACACGGCTGTCACACACACACACACACACACACACACACACACACAC 976

## ORIGIN

472 TCACACAGCTTTAAGATGACACACACACACACACACACACACACACAC 413

## ORIGIN

977 TGTCTAGCTGAAGCAATGACATGATGCTGTGTCTGTCTGTCTGTCTGT 1036

## ORIGIN

412 TCTCTATGTAAGCAATGATGCTGTGTCTGTGTCTGTGTGTGTGTGTGT 353

## ORIGIN

1037 CCTTCTGGAGATGCTGCTGATTAATTTTGTCTGTGTGTGTGTGTGTGT 1096

## ORIGIN

352 CCTTCTGGAGATGCTGCTGATTAATTTTGTCTGTGTGTGTGTGTGTGT 293

## ORIGIN

1097 TTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1156

## ORIGIN

232 TCCGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 233

## ORIGIN

1157 ATTTCGTCGTCATGCTGTCGTCATGCTGTCGTCATGCTGTCGTCATG 1216

## ORIGIN

232 ATTTCGTCGTCATGCTGTCGTCATGCTGTCGTCATGCTGTCGTCATG 173

## ORIGIN

1217 GTTTCGGCATTTATAGTCCCAACCTCCACGCTCTTCTAAGGAGGAG 1276

## ORIGIN

172 CCAC-----ACCTGCCAACCCACCTCCACAAACGCCAAAGATGCTATCA 122

## ORIGIN

1277 GGAACCTACCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 1330

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## ORIGIN

121 AGAAGAGTTTGTGACCGGCGCAAAAGATTGACAGCATGACAGCTGGCT 68

## JOURNAL COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1M11934 row: c column: 08  
High quality sequence stop: 756.

## FEATURES

source

1. 762

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:5367103"

/clone\_lib="NIH-MGC-94"

/tissue\_type="retina"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: eye; Vector: pCMV-Sport6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."

Note: this is a NIH-MGC Library."



[illegible]

Qy	883	ATGATGTCGTCCTCCCTGCGCCGTCGGGCTGGGCAATACACACCGTCATCATACACACC	942
Db	552	CCGACCGCGAAGTCGTCCACAGATCCCTCCGCGTATCTTCATGTCCTCAGCTTCCAGCTTGGCTCT	611
Qy	943	CAGAGCTCTGGCTCCCGGCGCTCTTTGGCTTAAGCTGTCTACGTGAAGGCAATCGACATC	10020
Db	612	GAGTGCACAAACCCCTTGGCGCGTGCAGCTTCCCAAGATTTCTATGTGAAGGCGCTTGATGTT	671
Qy	1003	TGATATGCTGTGTCTGCTCTTTGTGTCTTCCCTGCGCTTGCCTGAGATATGCTG	1054
Db	672	TGCGTTATGTGCTCTTCTCTTGGTGGTGTCTTCCCTGCGGAGATATGACAG	723
RESULT 13			
LOCUS	BI825190		
DEFINITION	BI825190	908 bp	mRNA
ACCESSION	603072067F1	NIH_MGC_119	Homo sapiens cDNA clone IMAGE:5164025 5',
KEYWORDS	NIH_MGC		mRNA sequence.
SOURCE	BI825190.1	GI:15936740	
ORGANISM	human.		
REFERENCE	Human sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
COMMENT	NIH-MGC http://mgc.ncl.nih.gov/		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs@mail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMNL at:		
	http://image.lncl.nih.gov		
	Plate: LMNL1407	row: e	column: 18
FEATURES	High quality sequence stop: 844.		
source	Location/Qualifiers		
	1..908		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5164025"		
	/clone_1b="NIH_MGC_119"		
	/tissue_type="medulla"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;		
	Site_2: EcoRV (destroyed); RNA source normal medulla from		
	anonymous male age 27. Library is oligo-dT primed and		
	directionally cloned (EcoRV site is destroyed upon		
	cloning). Average insert size 1.3 kb. Insert size range		
	0.9-3 kb. Library is normalized and enriched for		
	full-length clones and was constructed by C. Gruber		
	(Invitrogen). Research Genetics tracking code 013. Note:		
	this is a NIH MGC Library."		
BASE COUNT	341 a	156 c	190 g
ORIGIN			221 t
Query Match	10.4%:	Score 171.2;	DB 13; Length 908;
Best Local Similarity	74.4%:	Pred. No. 6e-26;	
Matches	229; Conservative	0; Mismatches	78; Indels
			1; Gaps
Qy	271	GACTACCGGGTGAATGTC-TTCTTGGCGGCAACAGTGAATACACACCGCTGTCTTACCG	329
Db	42	GATTACAGAGATATCTTCTTGGCGGCAAGAAATGAGATATCCCGCTTCGCTACAG	101
Qy	330	AGAAATATCCATGATCATCTCTGAGACCTTCGATCCCATGTGCGACCTATCTGGAACCC	389
Db	102	TGAATATCTGACGACTCTTTAGACCTTCGACCCCTCCATGTTGGACTCATTGGAACCC	161

390 AGACCTCTTCTTGTCTAATGAGAAAGGGCCACTTCATGAGTGACCGACACAA 449  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 162 TGATTTGTTCTTCCATGAAAGGGTGCCAACTTTCATGAGTGACCGACACAA 221  
450 GTTCTGGGCACTTCAGAAATGGGATGCTGTACAGCAAGCGGACCGCATTT 509  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 222 ATGTGTAAGAAATTTCAAAATGGAATGTTCTTATTCATTAATTAACCTTACACT 281  
510 GTCTCGCTGATGAGCACTTCAGAACTCCCATGAGATCCAGAGTCGACAGTCACT 569  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 282 TTCTGCTGCATGAGATTCAGAAATTTCCATGATGATCAAAACATGATTAATGCACT 341  
QY 570 TGAGAGCT 577  
11 | 11 | 11 |  
Db 342 GGAAGCT 349

RESULT 14  
AM280976 501 bp mRNA linear EST 26-JUL-2002  
LOCUS f148d02.y1 zebrafish adult brain Danio rerio cDNA 5' similar to  
DEFINITION gb:X52009 GLYCINE RECEPTOR ALPHA-1 CHAIN PRECURSOR (HUMAN);, mRNA  
SEQUENCE.  
AM280976  
ACCESSION AM280976.1 GI:6669525  
KEYWORDS EST.  
VERSION AM280976.1 GI:6669525  
TERMINUS  
SOURCE zebrafish.  
ORGANISM Danio rerio  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes  
; Cyprinidae; Danio.  
1 (bases 1 to 501)  
REFERENCE Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy,  
AUTHORS S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood,  
, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B.,  
Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,  
Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.  
and Wilson, R.  
TITLE WashU zebrafish EST Project 1998  
JOURNAL Unpublished (1998)  
COMMENT Other ESTs: f148d02.x1  
Contact: Stephen L. Johnson  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: zbrafish@watson.wustl.edu  
CDNA Library Preparation: John Ngai. CDNA Library Arrayed by:  
Matthew Clark. DNA Sequencing by: Washington University Genome  
Sequencing Center Clone distribution: Genome Systems, St. Louis,  
Missouri (web address: www.genomesystems.com) (email contact:  
info@genomesystems.com) and Research Genetics, Huntsville, Alabama  
(web address: www.resgen.com) (email contact: info@resgen.com) and  
Ressourcenzentrum Primatendatenbank, Berlin, Germany (web address:  
www.rzpd.de)  
Seq primer: T7  
High quality sequence stop: 328.  
FEATURES  
source  
1. 501  
Location/Qualifiers  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone\_lib="zebrafish adult brain"  
/sex="mixed male and female"  
/tissue\_type="brain"  
/dev\_stage="adult"  
/lab\_host="E. coli DH10B"  
/note="vector: pZIRPlox, site\_1: NotI, site\_2: SalI;  
original library was constructed in lambdaZIRPlox. Mass  
excision of the cDNA library was performed to yield  
pZIRPlox plasmids. Insert check was done in original  
library."  
BASE COUNT 114 a 137 c 123 g 127 t  
ORIGIN

Query Match 10.3%, Score 168.6; DB 10; Length 501;  
Best Local Similarity 76.4%; Pred. No. 2.4e-25;  
Matches 207; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

123 CATGCCCCCTCATATTCCTAGCAAACTATGAGGGGAGACATCTGATATGATGCAG 182  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 231 CATGCTCCCT 290  
183 GATTCGGCCCAATTTTAAAGGCCCAACCCGTGAGCTGACCTGCAACATCTTCATCAACAG 242  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 291 CATTGACCAAACTTCAAAAGGCTCTCTGACGTACGCTGATATATATTCATCAACAG 350  
243 TTTCAGCTCCGTACCAAGACACCAATGAGTACCGGTGATGTCTTCTGGCGGCAACA 302  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 351 CTTCGCTTCAGTCAAGAACTACTATGACTACAGGGTGAACATTAATCTCCGCGGCAAAA 410  
303 GTGGAATGACCCAGCCCTGCTCTACCGAGAAATATCTGATCTCTGACCTGATCC 362  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 411 GTGGAAGACACCTCTGACCTGCTCTACAGGAGATATCCGACTGCTTTGGATCTGAGACC 470  
QY 363 CTCATGCTGAGCTCTATCTGGAACCGAC 393  
11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 |  
Db 471 GTCCATGCTGAGCTCTGAGAACTGAGAACTGAC 501

RESULT 15  
B0938794 975 bp mRNA linear EST 21-AUG-2002  
LOCUS B0938794  
DEFINITION AGENCOURT 8930535 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:646531  
5' mRNA sequence.  
ACCESSION B0938794 GI:22354272  
VERSION B0938794.1 GI:22354272  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 975)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs.fremail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL13990 row: p column: 20  
High quality sequence stop: 638.  
FEATURES  
source  
1. 975  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone\_image:646531  
/clone\_lib="NIH\_MGC\_94"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 263 a 234 c 221 g 257 t  
ORIGIN  
Query Match 9.7%; Score 159.6; DB 14; Length 975;  
Best Local Similarity 54.3%; Pred. No. 1.6e-23;  
Matches 376; Conservative 0; Mismatches 274; Indels 42; Gaps 1;



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:14 ; Search time 62.6734 Seconds  
(without alignments)  
916.354 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVLPATLSFLLMTLPGQ.....POPPAPLREGETTRKLYVD 431

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2163	96.0	417	23	AAE21178 Human TRICH-22 pro
2	2157	95.8	417	21	AAE19336 Amino acid sequenc
3	871	38.7	496	23	ABE57052 Mouse ischemic co
4	754.5	33.5	517	22	AAE13035 R. sanguineus glut
5	742.5	33.0	450	22	AAE13037 R. sanguineus glut
6	742.5	33.0	450	22	AAE13038 R. sanguineus glut
7	737	32.7	444	23	AAE16395 Heliothis virescen
8	737	32.7	450	23	AAE16439 Heliothis virescen
9	736	32.7	537	23	AAE13312 Dermacentor variab
10	732	32.5	537	23	AAE13314 Dermacentor variab

11	731.5	32.5	455	20	AAW97860
12	731	32.5	537	23	AAE13313
13	730.5	32.4	456	17	AAW05246
14	729	32.4	606	22	ABE5121
15	726	32.2	637	22	AAE1077
16	725.5	32.2	462	22	AAE00863
17	725.5	32.2	462	22	AAE00864
18	724.5	32.2	459	20	AAE01636
19	724.5	32.2	467	20	AAE1634
20	724.5	32.2	481	20	AAE1635
21	724.5	32.2	496	20	AAE1633
22	720	32.0	488	19	AAW69285
23	720	32.0	488	19	AAW69286
24	720	32.0	606	23	AAE13315
25	720	32.0	637	21	AAE1074
26	720	32.0	637	21	AAE1075
27	719.5	31.9	454	22	ABE53334
28	719	31.9	397	22	AAE13034
29	718	31.9	637	21	AAE1076
30	717	31.8	496	17	AAE89336
31	717	31.8	496	20	AAE6161
32	716	31.8	561	19	AAW97413
33	715	31.7	496	22	ABE56256
34	712	31.6	474	14	AAE1188
35	712	31.6	474	15	AAE59866
36	704	31.3	637	14	AAE34035
37	698.5	31.0	478	22	AAE0084
38	692.5	30.8	474	22	AAE39098
39	692	30.7	473	23	ABE57078
40	690	30.6	513	22	AAE00861
41	690	30.6	513	22	AAE00862
42	678.5	30.1	452	17	AAE97299
43	674.5	30.0	487	23	AAE1289
44	673	29.9	487	17	AAE8361
45	670.5	29.8	510	17	AAE88360

## ALIGNMENTS

RESULT 1	AAE21178				
ID	AAE21178	standard; Protein: 417 AA.			
XX	XX				
AC	AAE21178:				
XX	XX				
DT	01-JUN-2002	(first entry)			
XX	XX				
DE	Human TRICH-22 protein.				
XX	XX				
KW	Human; transporter and ion channel; TRICH-22; transport disorder; angina;				Cat flea glutamate
KW	amniotic lateral sclerosis; cystic fibrosis; neuromuscular disorder;				Dermacentor variab
KW	cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;				Drosophila glutama
KW	depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;				Drosophila melanog
KW	cell proliferated disorder; infertility; arteriosclerosis; gene therapy;				D. stimulans GABA I
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;				Short form of S. a
KW	mysthenia gravis; multiple sclerosis; metabolic disorder; hypertension;				GABA-gated chlorid
KW	acquired immune deficiency syndrome; immunological disorder; scleroderma;				GABA-gated chlorid
KW	endocrine disorder; autoimmune thyroiditis; rheumatoid arthritis; goitre;				GABA-gated chlorid
KW	cardiac myopathy; amnesia; toxic myopathy; Addison's disease; infection;				GABA-gated chlorid
KW	epilepsy; mental disorder; myocarditis; Crohn's disease; Grave's disease;				Dermacentor melano
KW	muscle disorder; stroke; dementia; anxiety; AIDS; asthma; cirrhosis.				D. melanogaster po
XX	XX				D. melanogaster CA
OS	Homo sapiens.				Drosophila melanog
XX	XX				Dermacentor variab
FH	Key	Location/Qualifiers			D. melanogaster GA
FT	Peptide	1..28			Insect GABA-A recept
FT	Protein	29..417			Lucilia cupripes GA
FT	Protein	/label= Signal_peptide			Drosophila melanog
FT	Protein	/note= "Mature human TRICH-22 protein"			GABA-A receptor be
FT	Domain	44..341			Sequence of a GABA
FT	Domain	/note= "Neurotransmitter-gated ion channel domain"			Human polypeptide
FT	Domain	172..186			Human polypeptide

FT Domain /note="Neurotransmitter-gated ion channel motif"  
FT 255..279  
FT /note="Transmembrane domain"  
FT 320..339  
FT Domain /note="Transmembrane domain"  
XX WO200212340-A2.  
XX 14-FEB-2002.  
XX 01-AUG-2001; 2001WO-US24217.  
XX 03-AUG-2000; 2000US-223269P.  
XX 10-AUG-2000; 2000US-224456P.  
XX 18-AUG-2000; 2000US-226410P.  
XX 25-AUG-2000; 2000US-228140P.  
XX 31-AUG-2000; 2000US-228067P.  
XX 08-SEP-2000; 2000US-231434P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Yue H, Thornton M, Ramkumar J, Tang YT, Azimzai Y, Baughn MR;  
PI Yang J, Yao MG, Lal P, Walla NK, Gandhi AR, Hafalia AJA;  
PI Nguyen DB, Patterson C, Elliott VS, Tribouley CM, Lu DAM, Xu Y;  
PI Reddy R, Hernandez R, Borowsky ML, Lo TP, Lu Y, Policky JL;  
PI Greene BD, Sanjanwala MS, Raumann BE, Burford N, Ison CH, Lee EA;  
PI Ding L, Das D, Kallik DA, Khan FA, Sellhammer JJ;  
DR WPI: 2002-206330/26.  
DR N-PSDB: AAD33667.  
PT New human transporters and ion channels polypeptides and  
PT polynucleotides for diagnosing, preventing or treating transport,  
PT neurologicall, muscle, immunological and cell proliferative disorders  
PS Claim 66; Page 181-182; 230pp; English.

XX The invention relates to human transporter and ion channel polypeptides  
CC designated TRICH and nucleic acid molecules encoding such polypeptides.  
CC TRICH sequences are useful for diagnosis, treatment and prevention of  
CC transport, muscle, neurological, immunological and cell proliferative  
CC disorders. Transport disorders include akinesia, amytrophic lateral  
CC sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular  
CC dystrophy, diabetes mellitus, diabetes insipidus, myasthenia gravis,  
CC myocarditis, prostate cancer, cardiac disorders associated with  
CC transport e.g. polymyositis, bradyarrhythmia, dermatomyositis, angina,  
CC neurological disorders associated with transport e.g. amnesia, bipolar  
CC disorder, depression, Tourette's disorder, schizophrenia, other disorders  
CC associated with transport e.g. neurofibromatosis, sickle cell anaemia,  
CC Wilson's disease, cataracts, infertility, hypoglycaemia, hypoglycaemia,  
CC goitre, Cushing's disease, hypercholesterolaemia and cystinuria. Cell  
CC proliferated disorders include cancer, actinic keratosis, cirrhosis,  
CC arteriosclerosis, atherosclerosis, bursitis, hepatitis and psoriasis.  
CC Neurological disorders include Alzheimer's, Pick's and Parkinson's  
CC disease, amyotrophic lateral sclerosis, epilepsy, stroke, Huntington's  
CC disease, multiple sclerosis, dementia and other extrapyramidal disorder,  
CC motor neuron disorder, prion disease, metabolic disease of the nervous  
CC system and other developmental disorders of the central nervous system,  
CC neuromuscular disorders, metabolic, endocrine and toxic myopathies,  
CC periodic paralysis, mental disorders including mood, anxiety, and  
CC immunological disorders include acquired immune deficiency syndrome  
CC (AIDS), adult respiratory distress syndrome, Addison's disease,  
CC allergies, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, Crohn's disease, atopic dermatitis,  
CC Grave's disease, glomerulonephritis, rheumatoid arthritis, scleroderma,  
CC systemic lupus erythematosus, systemic sclerosis, ulcerative colitis,  
CC haemodialysis, uveitis, viral, bacterial, fungal, parasitic, protozoal,  
CC hematinic infections and trauma; and muscle disorders include cardiac  
CC myopathy, myocarditis, polymyositis, arrhythmias and hypertension. The  
CC TRICH polynucleotides are used in gene therapy. The present sequence  
CC is human TRICH-22 protein.  
XX Sequence 417 AA:  
SQ

Query Match 96.0%; Score 2163; DB 23; Length 417;  
Best Local Similarity 96.5%; Pred. 1.4e-216;  
Matches 416; Conservative 0; Mismatches 1; Indels 14; Gaps 1;  
QY 1 MTTLVPAATLSELLMTLPGVLLRVALAKEEVKSGTSGSPSPDLKLMGRTSGYDA 60  
DB 1 MTTLVPAATLSELLMTLPGVLLRVALAKEEVKSGTSGSPSPDLKLMGRTSGYDA 60  
QY 61 RIRPFKGPVNVNCFINFSSTYKTTMYRVNVPFRQWNPRLSYREYPPDSLDLD 120  
DB 61 RIRPFKGPVNVNCFINFSSTYKTTMYRVNVPFRQWNPRLSYREYPPDSLDLD 120  
QY 121 PSMDSIKRPLDFEANEKANFHEVTTNKLIRFKNGNVYSIRLTILISCLMDLKNFP 180  
DB 121 PSMDSIKRPLDFEANEKANFHEVTTNKLIRFKNGNVYSIRLTILISCLMDLKNFP 180  
QY 181 MDIQTCTMOLESSSILSPLSLSVGYTKMDLFEWLEDAPAVQVAEGLTLPOFILRD 240  
DB 181 MDIQTCTMOLESSSILSPLSLSVGYTKMDLFEWLEDAPAVQVAEGLTLPOFILRD 240  
QY 241 EKDLACCTKHNNTGKFTGIEVKFHLEROMGYLLQMTIPSLIYILSVSWFINMDAPA 300  
DB 227 EKDLACCTKHNNTGKFTGIEVKFHLEROMGYLLQMTIPSLIYILSVSWFINMDAPA 286  
QY 301 RVGLGITTVLMTTQSSGSRASLPKVSYSKADIDIMNAVCLFEVPAALLEVAINFVSROH 360  
DB 287 RVGLGITTVLMTTQSSGSRASLPKVSYSKADIDIMNAVCLFEVPAALLEVAINFVSROH 346  
QY 361 KEFIRLRORRORLEEDIIODESREYFPGYGLGHCLQARDGPMGSGTYSPPAPPLR 420  
DB 347 KEFIRLRORRORLEEDIIODESREYFPGYGLGHCLQARDGPMGSGTYSPPAPPLR 406  
QY 421 EGETTRKLYVD 431  
DB 407 EGETTRKLYVD 417

RESULT 2  
AAB19336  
ID AAB19336 standard; Protein; 417 AA.  
XX AAB19336;  
AC 19-FEB-2001 (first entry)  
XX 19-FEB-2001 (first entry)  
DT Amino acid sequence of a human ataxia protein.  
XX Human; ataxia; gene therapy.  
KM Homo sapiens.  
OS WO200058461-A1.  
XX 05-OCT-2000.  
PD 23-MAR-2000; 2000WO-EP02600.  
XX 26-MAR-1999; 99EP-0106343.  
PR (RAPP/) RAPPOLD-HOERBRAND G.  
XX RAPPOLD-HOERBRAND G;  
PI WPI: 2000-656166/63.  
XX N-PSDB: AAC61678.  
DR Novel nucleic acid sequence encoding human ataxia protein for screening  
XX compounds useful for treating disorders relating to mutations in ataxia  
PT gene  
XX Claim 10; Page 20-21; 47pp; English.  
PS

CC The present sequence represents a human ataxia protein. The ataxia  
 CC protein and polynucleotides are useful for diagnosing and treating  
 CC disorders related to ataxia. Ataxia gene sequences are useful in  
 CC gene therapy, and as diagnostic tools or reagents for identifying and  
 CC characterizing genetic defect involved in the disorders and diseases  
 CC related to ataxia.

XX Sequence 417 AA;

Query Match 95.8%; Score 2157; DB 21; Length 417;  
 Best Local Similarity 96.1%; Pred. No. 3,7e-216;  
 Matches 414; Conservative 1; Mismatches 2; Indels 14; Gaps 1;

QY 1 MTLVATISFLMLTLPQVLLRVALAKEEVKSGTKSGQSPSPDFDKMGRTSGYDA 60  
 DB 1 MTLVATISFLMLTLPQVLLRVALAKEEVKSGTKSGQSPSPDFDKMGRTSGYDA 60  
 QY 61 RIRPNFKGPPVNTVCNIFINSFSSVTKTMDYRVNVLROQNDPRLSYREYDDSLDD 120  
 DB 61 RIRPNFKGPPVNTVCNIFINSFSSVTKTMDYRVNVLROQNDPRLSYREYDDSLDD 120  
 QY 121 PSMUDSIMKPDLEFANEKANFHEVTTDNKLRIFKNGNVLYSIRLTLLISCLMDLKNP 180  
 DB 121 PSMUDSIMKPDLEFANEKANFHEVTTDNKLRIFKNGNVLYSIRLTLLISCLMDLKNP 180  
 QY 181 MDIOTCTMOLESSILCSPLPSLSVGYTMDLVEFEMLEDAVAVOAEGLTPPFLRD 240  
 DB 181 MDIOTCTMOLE-----RFGYTKMDLVEFEMLEDAVAVOAEGLTPPFLRD 226 33  
 QY 241 EKDLGCTTKHYNTGKFTCEVFKFLEBROMGYLIQWIPSLIYLISWSFWINMDAPA 300  
 DB 241 EKDLGCTTKHYNTGKFTCEVFKFLEBROMGYLIQWIPSLIYLISWSFWINMDAPA 300  
 QY 227 EKDLGCTTKHYNTGKFTCEVFKFLEBROMGYLIQWIPSLIYLISWSFWINMDAPA 286  
 DB 301 RVLGTTTTLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYFAALLEYAAINPVSROH 360  
 QY 287 RVLGTTTTLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFYFAALLEYAAINPVSROH 346  
 QY 361 KEFIRLRRORRORLEEDIQESRFYFGYGLCHCLQARDGPMESGSGISPOPPAPLLR 420  
 DB 347 KEFIRLRRORRORLEEDIQESRFYFGYGLCHCLQARDGPMESGSGISPOPPAPLLR 406  
 QY 421 EGETTRKLYVD 431  
 DB 407 EGETTRKLYVD 417

RESULT 3

ABB57052 standard; Protein: 496 AA.

XX ABB57052;

DT 07-MAR-2002 (first entry)

DE Mouse Ischaemic condition related protein sequence SEQ ID NO:86.

XX Mouse: Ischaemia; compressive ischaemia; occlusive ischaemia;  
 KM vasospastic ischaemia; ischaemic condition; ischaemic disease.

XX Mus musculus.

XX MO200188188-A2.

XX 22-NOV-2001.

XX 18-MAY-2001; 2001MO-JP04192.

XX 18-MAY-2000; 2000JP-0145977.

XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.

XX Ishikawa K, Asei S, Takahashi Y, Nagata T, Ishii Y;

XX

DR WPI; 2002-034733/04.

DR N-PSDB; AB199254.

PT Examining the ischemic condition (e.g. occlusive ischaemia) by measuring

PT expression levels of particular genes defined in the specification or

PT genes -

PS Claim 2; Page 269-271; 2690pp; English.

XX The present invention describes a method for examining ischaemic  
 CC conditions, comprising measuring the expression levels of particular  
 CC genes (I) in a test sample or determining the expression profile of a  
 CC gene group in the sample comprising genes selected from (I). The method  
 CC is useful for examining the ischaemic condition (e.g. compressive  
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
 CC expression levels of particular genes (AB199202 to AB199912, encoding  
 CC the protein sequences in ABB57020 to ABB57374) or by determining the  
 CC expression profile of a gene group comprising these genes. The  
 CC expression levels or expression profiles produced by these genes are  
 CC used as an indicator when screening for ischaemic condition-improving  
 CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914  
 CC represent PCR primers for a mouse ischaemic condition related sequence,  
 CC which are used in the exemplification of the present invention.

XX Sequence 496 AA;

Query Match 38.7%; Score 871; DB 23; Length 496;

Best Local Similarity 48.5%; Pred. No. 1e-81;  
 Matches 188; Conservative 61; Mismatches 87; Indels 52; Gaps 11;

QY 9 LSFLLMTLPQVLLRVALAKEEVKSGTKSGQ-----PMSPDFDK 50  
 DB 7 ISFILLMSL-----LPEDACKKESKKKQKQYLCPSQSPEDLARVPNSTSLNR 62  
 QY 51 LMGRTGYDARIRPNFKGPPVNTVCNIFINSFSSVTKTMDYRVNVLROQNDPRL--- 107  
 DB 63 LL--VSYDRIIRPNFKGPPVNTVCNIFINSFSSVTKTMDYRVNVLROQNDPRLKLP 119  
 QY 108 -SYREYDDSLDDPSMUDSIMKPDLEFANEKANFHEVTTDNKLRIFKNGNVLYSIRL 166  
 DB 120 SDFR--GSDALTYDPTWKCLMKPDLFEFANEKANFHEVTTDNKLRIFKNGNVLYSIRL 177  
 QY 167 TLLISCLMDLKNPMDIOTCTMOLESSILCSPLPSLSVGYTMDLVEFEMLEDAVAVO 226  
 DB 178 SITLSCPLDLTLPMDIQCKMQLE-----SFGYTTDBLRIMQSGDP-VQ 222  
 QY 227 VAEGLTLPQ-ILRDEKDLGCTTKHY-NTGKFTCEVFKFLEBROMGYLIQWIPSLIYL 284  
 DB 223 L-EKIALPQFDIKKEDIYGNCTKYKGYTCVEVIFTLRQVGYMMGVAPPLLIY 281  
 QY 285 ILSWVSPINMDAAPAVGIGITTYLTMTQSSGRASLPKVSYYKAIIDIMAVCLLFVF 344  
 DB 282 VLSWLSFWINPDASAARVPPIGIFSULASACTTLAELPKVSYYKALDWLACLLFGF 341  
 QY 345 AALLEYAAINPVSROHKEFIRLRRORR 372  
 DB 342 ASLVEYAVVQVMUNPK---RVEAEKRR 366

RESULT 4

AAE13039 standard; Protein: 517 AA.

XX AAE13039;

XX 28-JAN-2002 (first entry)

DE R. sanguineus glutamate-gated chloride channel 1 protein, T32.

XX Brown dog tick; glutamate-gated chloride channel; GluCl1;

KW GluCl2; crop protection; insecticide; nematocide; acaricide;

KW T32 protein.

[illegible]

```

QY      390 YGLGHCLOARDGGPMESGIYS---PQPPAP 417
          |      | : | : | | | |
Db      410 VSSG-----EPACLMARTWPPPLP 429

```

RESULT 5  
AAE13037

AC. AAE13037;

DT 28-JAN-2002 (first entry)

DE R. sanguineus glutamate-gated chloride channel 1 protein, 112.

kw brown dog tick; gluamate-gated chloride channel; glucr;  
gluc12; crop protection; insecticide; acaricide;  
kw gluc12; crop protection; nematocide; acaricide;

XX

22 XX

XX 11-OCT-2001  
PD

AA  
PF 28-MAR-2001: 2001WO-11509905

PR 31-MAR-2000; 2000US-193934P

PA (MERI ) MERCK & CO INC.

PI Warmke JW, Yang Y, Cully DF, Hamelin MJ, et al. 2010. The

DR WPI; 2001-662963/76.  
DP N-PDDB; 8AD31305

Novel L-glutamate-gated chloride channels from *Rhizopneustes*

PT proteins, which are useful as insecticides, anthelmintics and

XX

XX

CC L-glutamate-gated chloride channel proteins (Gluc11 and Gluc12)

CC proteins are useful for identifying modulators. The compounds  
CC identified as modulators are useful for insecticidal or mitacidal

and/or nematocidal treatment for use in animal and human health and/or crop protection. The compounds are also useful in screening

CC species relevant to animal and human health, including worms,

CC functional Gluc11 and Gluc12 channel forms are useful

CC GluCl channel modulators. The present sequence is *R. sanguineus*  
CC GluCl1 protein, M13

Sequence 450 AA:

Query Match 33.08; Score 742.5; DB 22; Length 450;

Matches 157; Conservative 61; Mismatches 110; Indels 29; Gaps 8;

QY 48 LDKLMGRTSGYDARIRP---NEKGPVNVTCNIFINSESSVTKTTMDYRVNVFELRQOWN 103

Db 42 LDSIIGQ-GRYDCRIRPMGINNTDGPAL-VRVNI FVRSIGRIDDVMTMEYTVQMTFREQWR 99

QY 104 DPRLSYREYPDDSLDLDPSMLDSIWKPDLEFANEKGANFHEVTTDNKLLRIFKNGNVLYS 163

Db 100 DERLOYD LGOVRYLTTEPDKLWKPDLFESNEKEGHFNITMPNVLRIBNGDVLFS 159







```
XX 12-FEB-2002 (first entry)
DT Dermacentor variabilis clone 8 GABA-gated chloride channel protein.
XX
DE Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;
KW GABA-gated chloride channel; recombinant expression; domestic animal.
XX
OS Dermacentor variabilis.
XX
PN WO200174884-A1.
XX
PD 11-OCT-2001.
XX
PE 28-MAR-2001; 2001WO-US09955.
XX
PR 31-MAR-2000; 2000US-193791P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Zheng Y, Cully D, Ludmerer S;
XX
DR WPI: 2002-010778/01.
DR N-PSDB: AAD22070.
XX
XX New polypeptide useful for preventing or treating tick infestation. In
PT humans, dogs, cattle, horses, deer, or other wild or domesticated
PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid
PT (GABA)-gated chloride channel.
XX
PS Claim 1: Fig 1: 59pp; English.
XX
XX The invention relates to gamma-aminobutyric acid (GABA)-gated chloride
CC channels and their corresponding nucleic acid molecules. GABA-gated
CC chloride channel proteins and DNA's are useful for preventing and
CC treating tick infestation, particularly in humans, dogs, cattle, horses,
CC deer, or other wild or domesticated animals. The nucleic acids are useful
CC as hybridisation probes or Polymerase Chain Reaction primers for
CC identifying the presence of Dermacentor variabilis GABA-gated chloride
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic
CC acids are also useful for the recombinant expression of D. variabilis
CC GABA-gated chloride channel proteins. GABA-gated chloride channel
CC proteins exert toxic effects on other ticks or related parasites such as
CC mites. The present sequence is Dermacentor variabilis clone 8 GABA-gated
CC chloride channel protein.
XX
SO Sequence 537 AA:
Query Match 32.7%; Score 736; DB 23; Length 537;
Best Local Similarity 40.9%; Pred. No. 1.4e-67;
Matches 158; Conservative 68; Mismatches 110; Indels 50; Gaps 9;
OY 55 TSGYDARIRPNFKGPVAVNTCNIFNSFSYTKTMDYRVAVFLRQONDRSLSYREYD 114
DB 52 TRGYDRRVRPNYGVGVETVMTQIISIVSEVOMDPTSFYRQSWDRSLSFQKSPD 111
OY 115 -DSLDLPMSMDSITWKPDLPFANEKANFHEVTTDNKLRIFKNGNVLSIRLTLISCL 173
DB 112 LESMTVGAEVAERLWVPDPTFANEKSAFHAATTPNFTLRIGSGGEVFRSIRLTVTASCP 171
OY 174 MDLKNFMDIQTCTMQLSESSILCSPLPSLSLVGYTKMDLVFEWLEDAPAVOVAEGTL 233
DB 172 MDLRFPMDDROACITIEI-----SFGYTKMDIRYMSDDGTSVRIAKEVEL 217
OY 234 PQF-----ILRDEKDLGCTKRYNNGKFCIEVKFHELEQOMGYLLIOMTIPSLIIVISW 289
DB 218 PQFVGLGHVQAKAEALITGNYS--RLVC-EIRF--ARSMGYLLIQTIPAGLIVISW 272
OY 290 SFWIMMDAAPAVGIGITVTLMTTOSGSRASLPKSYKAIIDMMVAVCLLEFAALE 349
DB 273 SFWILHRODSPARVALGVTTLMTTLMSTNAALPKIYSYKSIDVYLTGTCVVMFTALLE 332
OY 350 YAAINVSROHKEFLRLRRORRORLEEDIIQESHFYRGYGLGCLQARDGGMWEGSGI 409
```

```
DB 333 YAAVGYLCK-RITMRKTRCQQLAKLAHQHQR-----RCAASSNERSSEPL 378
OY 410 YSPQ-----PAPLRGE 423
DB 379 ASPEVSIVKTVGSCVCPAAVASOGO 404
RESULT 10
AEI3314
ID AEI3314 standard; Protein: 537 AA.
XX
AC AEI3314:
XX
XX 12-FEB-2002 (first entry)
DT Dermacentor variabilis clone 5 GABA-gated chloride channel protein.
XX
DE Gamma-aminobutyric acid; GABA; tick infestation; mite; antiparasitic;
KW GABA-gated chloride channel; recombinant expression; domestic animal.
XX
OS Dermacentor variabilis.
XX
PN WO200174884-A1.
XX
PD 11-OCT-2001.
XX
PE 28-MAR-2001; 2001WO-US09955.
XX
PR 31-MAR-2000; 2000US-193791P.
XX
PA (MERI ) MERCK & CO INC.
XX
PI Zheng Y, Cully D, Ludmerer S;
XX
DR WPI: 2002-010778/01.
DR N-PSDB: AAD22072.
XX
XX New polypeptide useful for preventing or treating tick infestation. In
PT humans, dogs, cattle, horses, deer, or other wild or domesticated
PT animals, comprises the Dermacentor variabilis gamma-aminobutyric acid
PT (GABA)-gated chloride channel.
XX
PS Claim 1: Fig 1: 59pp; English.
XX
XX The invention relates to gamma-aminobutyric acid (GABA)-gated chloride
CC channels and their corresponding nucleic acid molecules. GABA-gated
CC chloride channel proteins and DNA's are useful for preventing and
CC treating tick infestation, particularly in humans, dogs, cattle, horses,
CC deer, or other wild or domesticated animals. The nucleic acids are useful
CC as hybridisation probes or Polymerase Chain Reaction primers for
CC identifying the presence of Dermacentor variabilis GABA-gated chloride
CC channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic
CC acids are also useful for the recombinant expression of D. variabilis
CC GABA-gated chloride channel proteins. GABA-gated chloride channel
CC proteins exert toxic effects on other ticks or related parasites such as
CC mites. The present sequence is Dermacentor variabilis clone 5 GABA-gated
CC chloride channel protein.
XX
SO Sequence 537 AA:
Query Match 32.5%; Score 732; DB 23; Length 537;
Best Local Similarity 40.7%; Pred. No. 3.6e-67;
Matches 157; Conservative 68; Mismatches 111; Indels 50; Gaps 9;
OY 55 TSGYDARIRPNFKGPVAVNTCNIFNSFSYTKTMDYRVAVFLRQONDRSLSYREYD 114
DB 52 TRGYDRRVRPNYGVGVETVMTQIISIVSEVOMDPTSFYRQSWDRSLSFQKSPD 111
OY 115 -DSLDLPMSMDSITWKPDLPFANEKANFHEVTTDNKLRIFKNGNVLSIRLTLISCL 173
DB 112 LESMTVGAEVAERLWVPDPTFANEKSAFHAATTPNFTLRIGSGGEVFRSIRLTVTACGP 171
```



XX 28-MAR-2001; 2001WO-US09955.  
PF 31-MAR-2000; 2000US-193791P.  
XX (MERI ) MERCK & CO INC.  
XX Zheng Y, Cully D, Ludmerer S;  
XX WPI: 2002-010778/01.  
DR N-PSDB; AAD22071.  
XX New polypeptide useful for preventing or treating tick infestation, in  
PT humans, dogs, cattle, horses, deer, or other wild or domesticated  
PT animals, comprises the Dermancentor variabilis gamma-aminobutyric acid  
PT (GABA)-gated chloride channel  
XX  
XX Claim 1; Fig 1; 59pp; English.  
XX  
XX The invention relates to gamma-aminobutyric acid (GABA)-gated chloride  
XX channels and their corresponding nucleic acid molecules. GABA-gated  
XX chloride channel proteins and DNA's are useful for preventing and  
XX treating tick infestation, particularly in humans, dogs, cattle, horses,  
XX deer, or other wild or domesticated animals. The nucleic acids are useful  
XX as hybridisation probes or Polymerase Chain Reaction primers for  
XX identifying the presence of Dermancentor variabilis GABA-gated chloride  
XX channel nucleic acid or DNA's encoding for a GABA receptor. The nucleic  
XX acids are also useful for the recombinant expression of D. variabilis  
XX GABA-gated chloride channel proteins. GABA-gated chloride channel  
XX proteins exert toxic effects on other ticks or related parasites such as  
XX mites. The present sequence is Dermancentor variabilis clone 9 GABA-gated  
XX chloride channel protein.  
SO Sequence 537 AA:  
Query Match 32.5%; Score 731; DB 23; Length 537;  
Best Local Similarity 40.7%; Pred. No. 4.6e-67;  
Matches 157; Conservative 69; Mismatches 110; Indels 50; Gaps 9;  
QY 55 TSGYDARIRPNFGPVPVNTCNIFINSFSSVTKTTMDYRVNVLROQWNPRLSYREYD 114  
DB 52 TGTGDRVRPNFTGGVEVGVMTQIISIVSVSDVDFRQSWMRERLSFQKSPD 111  
QY 115 -DSLDLDPMSLSTINPKDLFFANKEGANFHEVTTDNKLRIFKGNVLSIRLTLSCL 173  
DB 112 LBSMTVGAEVARIWPDFFANESKAYFHAATTPMFLRISGGEVFRSIRLTIVASCP 171  
QY 174 MDLKNFMDIQTCTMOLESSSLCSPLPSLSVGYTMKDLVEFWLEDAPAVOAGLTL 233  
DB 172 MDLRYPMDROACTIEIE-----SFGYTMKDRIYRMSDGDTSVRIAKEVEL 217  
QY 234 POF-----ILRDEKDLCCCKHYNTGKFTCTEVEKFLHEROMGYLLIOMYIPSLILVLSW 289  
DB 218 POKVUGHVOKAKEVLLTGNYS--RLVC-EIRF--ARSGYLLIYIPAGLIVISWV 272  
QY 290 SFWINDDAARVAVGLITTVLTMTTOSGSRASLPSVSYKAIDIMAAVCLLFVFAALLE 349  
DB 273 SFPLHNASARVALGVTVLTMTTILMSSINAPLPISTVSKSIDVYLGTGCFVAVFALLE 332  
QY 350 YAAINVSROHKEFFILRRRORRORLEDDIIQESRFFRGYGLGHLQIARDGSPMEGSGT 409  
DB 333 YAAVGYLGR--RITMKRTROOLAKLAEOHRQ-----RCAAAASNPSSEPLL 378  
QY 410 YSPQ-----PAPLREGE 423  
DB 379 ASPEVSIKTVGSCVCPAAVASOQ 404  
RESULT 13  
AAW05246  
ID AAW05246 standard; protein: 456 AA.  
XX  
AC AAW05246;

XX 09-FEB-1997 (first entry)  
DT Drosophila glutamate-gated chloride channel.  
XX  
XX DE glutamate gated chloride channel; GluCl; ligand gated channel;  
XX KW avermectin binding protein; glutamate binding protein;  
XX KM antiparasitic; anthelmintic; helminthiasis; ectoparasiticide;  
XX acaricide; insecticide; pesticide.  
XX Drosophila melanogaster strain Oregon R.  
XX  
XX Key Location/Qualifiers  
FH Peptide 1..22  
FT /label= Sig-peptide  
FT Protein 23..456  
FT /label= Mat\_protein  
XX  
XX WO9634940-A1.  
XX  
XX 07-NOV-1996.  
XX  
XX 01-MAY-1996; 96WO-US06035.  
XX  
XX 05-MAY-1995; 95US-0435933.  
XX (MERI ) MERCK & CO INC.  
XX  
XX Arena JP, Cully DF, Liu KK, Pareess PS;  
XX WPI: 1996-506147/50.  
XX N-PSDB; AAT43596.  
XX  
XX New nucleic acid encoding Drosophila avermectin and/or glutamate  
PT binding protein - useful for isolating cpds. used as antiparasitic  
PT agents for the treatment or prevention of helminthiasis in domestic  
PT animals  
XX  
XX Claim 11; Page 45-46; 64pp; English.  
XX  
XX A novel Drosophila glutamate gated chloride channel (GluCl) (AAW05246)  
XX is selectively opened by either avermectin or glutamate and  
XX represents a target of avermectin action in arthropods. Its amino  
XX acid sequence was deduced from a Drosophila head cDNA clone (AAT43596).  
XX Recombinant GluCl, or host cells expressing it, can be used to screen  
XX for modulators of GluCl that can be used as ectoparasiticide,  
XX antiparasitic, anthelmintic, acaricide or insecticide agents for  
XX animals or plants.  
XX  
XX Sequence 456 AA:  
Query Match 32.4%; Score 730.5; DB 17; Length 456;  
Best Local Similarity 42.3%; Pred. No. 4e-67;  
Matches 154; Conservative 61; Mismatches 120; Indels 29; Gaps 7;  
QY 21 VLLRVALAKEEVSCKTGSGPMSDFLDKLMGRTSGYDARIRP---NFKGPVNTCN 77  
DB 10 ILTFASLCSLSLANNKAVNREKEKYLVDIIG-AGKYDARIRPSGNGDGAIRYNL 68  
QY 78 FINSFSSVTKTTMDYRVNVLROQWNPRLSYREYDPSDLDPMSLSTINPKDLFFANE 137  
DB 69 FVRSIMTISDIKMEYSVQLTFRFQWMDERLKFDDIOGRLLKYLTLTEANRVMPDLPFSNE 128  
QY 138 KGANFHEVTTDNKLRIFKGNVLSIRLTLSCLMDLKNFMDIQTCTMOLESSSLC 197  
DB 129 KEGFPHNITMPNYYIRIFPNGSVLSYRISLTACPNKLKLYPLDRQICSLRW----- 181  
QY 198 SPLPSLSVGYTMKDLVEFWLEDAPAVOAGLTLPOFILLRDEKDL-GCCTHYNTGKF 256  
DB 182 -----ASYGMVTNDLVFLMKRGDP-VGVYKKLHLRPLF--EKFLTDVCNSKTYTGEY 231  
QY 257 TCIEVKFLHEROMGYLLIOMYIPSLILVLSWSPWINDDAARVAVGLITTVLTMTTOS 316



CC beneficial insect may render the insect resistant to a pesticide and can  
CC be used in combination with pesticides in the field to reduce or  
CC eliminate the presence of harmful insects only. The insect GABA receptor  
CC shows critical pharmacological differences from the vertebrate receptor  
CC which may result in the development of insect-specific insecticides with  
CC greater safety for vertebrate exposure. This sequence represents a  
CC variant *Drosophila simulans* (fruitfly) GABA receptor R1MD-R described  
CC in the method of the invention.  
XX

SQ Sequence 637 AA:

Query Match 32.28; Score 726; DB 21; Length 637;

Best Local Similarity 36.7%; Pred. No. 1.9e-66;

Matches 156; Conservative 86; Mismatches 143; Indels 40; Gaps 9;

```
QY 5 VPATLSFLLLMTLPGOVLNVALAKE-----VKSGTGSGSPMSPSDFLKMGRTS 56
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 46 LPRT-PLTTW-----LAINMALIAQETGHRIRHTVQAATGGGMLGDVNIATLDSFSV 99
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 GYDARIRPNFGPPVNTCNIFINSFSSVTKTMDYRVNVELRQWNPRLSYREYPD-D 115
   || | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 100 SYDKRVNPNGGPPVEGVMTMYVLSVSEVLMDFLDFYERQFWTDPRLAYRKRPVE 159
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 116 SLDLPSMLDSIMKPDLEFANEGANPEVTTDNKLLRIFKNGNVLSIRLTLLSCIMD 175
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 160 TLSVGSEFIKNIWPDFFVNEKOSYFIATTSNEFIRVHSGSITRSIRLTITASCPMN 219
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 176 LKNFPMIOICTQMOLESSILCSPLSLSVGTMKDLVFEULEDAPANQVAGLTLPQ 235
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 220 LQFFPMRQLCHEIE-----SFGYTRMDIRYFMRDGLSVGMSSEVELPQ 265
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 236 FILRDEKDLCCCFKHYNTGKFCIEVKFHLEROMGYLLIOMYIPSLIIVILSWSPWIM 295
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 266 FRVLGHNQ-RATEINLTGNSRLACIQEVRSMGYLLIQIYIPSGLIIVISWVFWLNR 324
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 296 DAAPARVGLITTVLIMTTQSSGSRASLPVSVYKAIIDIMMAVCLLFVFAALLEYAINF 355
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 325 NATPARVGLGVTVLTMTLMSSTNALPKISYKSIDVYLGTCFVWVFASLLEYATVGY 384
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 356 VSRQHKFEIILRRQRQRLEEDIQESRYFRGYGLGHC---LQARDGPMESGIYSP 412
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 385 MAK-----RIOMRKORFMAIQIAEQKKOOLDGANOQOANPNPNANVGGP-GGVGVGPG 437
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 413 QPPAP 417
   | |
Db 438 GPGGP 442
```

Search completed: June 25, 2003, 17:15:32  
Job time : 64.6734 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:15:40 ; Search time 35.6756 Seconds  
(without alignments)  
1307.255 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPAFLSLMTLPGO.....PPAPALREGETTRKLYVD 431

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 41779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/BCI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2252	100.0	431	US-10-075-846-4	Sequence 4, Appl1
2	2163	96.0	417	US-10-075-846-2	Sequence 2, Appl1
3	1682.5	74.7	452	US-10-075-846-13	Sequence 12, Appl1
4	1678	74.5	337	US-10-075-846-12	Sequence 13, Appl1
5	1638	72.7	312	US-10-075-846-15	Sequence 15, Appl1
6	1602.5	71.2	465	US-10-075-846-11	Sequence 11, Appl1
7	1580.5	70.2	449	US-10-075-846-10	Sequence 10, Appl1
8	1549	68.8	298	US-09-969-844-14	Sequence 14, Appl1
9	737	32.7	450	US-09-969-844-14	Sequence 14, Appl1
10	724	32.1	533	US-09-808-602-87	Sequence 87, Appl1
11	724	32.1	533	US-09-808-198-74	Sequence 74, Appl1
12	719	31.9	397	US-10-239-420-2	Sequence 2, Appl1
13	678.5	30.1	452	US-10-211-673-12	Sequence 12, Appl1
14	655	29.1	422	US-10-239-420-5	Sequence 5, Appl1
15	652.5	29.0	440	US-09-510-662A-39	Sequence 39, Appl1
16	652.5	29.0	440	US-09-778-320-39	Sequence 39, Appl1
17	652.5	29.0	440	US-09-910-689-39	Sequence 39, Appl1
18	652.5	29.0	440	US-10-010-742-39	Sequence 39, Appl1
19	652.5	29.0	461	US-10-106-698-6308	Sequence 6308, Appl1

20	632.5	28.1	397	US-09-808-483-6	Sequence 6, Appl1
21	632.5	28.1	422	US-09-808-483-4	Sequence 4, Appl1
22	632.5	28.1	423	US-09-808-483-8	Sequence 8, Appl1
23	620	27.5	465	US-08-818-657-2	Sequence 2, Appl1
24	620	27.5	465	US-09-893-321-2	Sequence 2, Appl1
25	619.5	27.5	554	US-10-211-673-8	Sequence 8, Appl1
26	618	27.4	465	US-09-818-657-4	Sequence 4, Appl1
27	615	27.3	464	US-09-839-446-35	Sequence 35, Appl1
28	615	27.3	464	US-09-839-446-36	Sequence 36, Appl1
29	615	27.3	464	US-09-839-446-37	Sequence 37, Appl1
30	615	27.3	464	US-09-839-446-38	Sequence 38, Appl1
31	615	27.3	464	US-09-742-311-4	Sequence 4, Appl1
32	615	27.3	464	US-09-898-570-35	Sequence 35, Appl1
33	615	27.3	464	US-09-898-570-36	Sequence 36, Appl1
34	615	27.3	464	US-09-898-570-37	Sequence 37, Appl1
35	615	27.3	464	US-09-898-570-38	Sequence 38, Appl1
36	615	27.3	464	US-09-808-483-12	Sequence 12, Appl1
37	615	27.3	535	US-09-808-483-10	Sequence 10, Appl1
38	608	27.0	467	US-09-765-069-2	Sequence 2, Appl1
39	608	27.0	467	US-09-742-311-2	Sequence 2, Appl1
40	607	27.0	420	US-09-765-069-8	Sequence 8, Appl1
41	603	26.8	392	US-09-765-069-4	Sequence 4, Appl1
42	602	26.7	345	US-09-765-069-10	Sequence 10, Appl1
43	594.5	26.4	468	US-09-839-446-6	Sequence 6, Appl1
44	594.5	26.4	468	US-09-898-570-6	Sequence 6, Appl1
45	592.5	26.3	466	US-09-839-446-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1  
US-10-075-846-4  
; Sequence 4, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075, 846  
; PRIOR FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269, 535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 431  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-075-846-4

Query Match 100.0%: Score 2252; DB 9; Length 431;  
Best Local Similarity 100.0%: Pred. No. 1.1e-206;  
Matches 431: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTLVPAFLSLMTLPGOVLRLVALAKEVKSQTSOPMSDFDLKMGRTSGYDA	60
DB	1	MTLVPAFLSLMTLPGOVLRLVALAKEVKSQTSOPMSDFDLKMGRTSGYDA	60
QY	61	RIRNFKGPPNVNVCNIFINFSSTYKTTMDYRVNVLFRQOMNDRLSYREYPPDSLDLD	120
DB	61	RIRNFKGPPNVNVCNIFINFSSTYKTTMDYRVNVLFRQOMNDRLSYREYPPDSLDLD	120
QY	121	PSMLDSIKPDLFPAKEGAFHEVYTDNKLIRFKNGNVLSIRLTILSLMDLNKP	180
DB	121	PSMLDSIKPDLFPAKEGAFHEVYTDNKLIRFKNGNVLSIRLTILSLMDLNKP	180
QY	181	MDIOTCTMQLSSSTLSPLSLISVGYTKKDLVFEMLDPAVNOVAEGTLTLOFTLRD	240
DB	181	MDIOTCTMQLSSSTLSPLSLISVGYTKKDLVFEMLDPAVNOVAEGTLTLOFTLRD	240
QY	241	EKDLGCTCTKHVNTGKFTCIEVKFHLEROMGYLLIOMYIPSLIIVLSVSWFMINDAPA	300

```
Db 241 EKDLGCTKHYNTGKFTCLEVGFHLEKRGYLLIOMYIPSLIIVLISWSEFIMNDAAPA 300
QY 301 RVGLGITTIVLMTTQSSGSRASLPVSYKAIIDIMAVCLLFVFAALLEYAAINVSROH 360
Db 301 RVGLGITTIVLMTTQSSGSRASLPVSYKAIIDIMAVCLLFVFAALLEYAAINVSROH 360
QY 361 KEFIRLRQRQRORLEEDIOESRFYRGYGLGHCLQARDGSGMGSIGTSPQAPAPLIR 420
Db 361 KEFIRLRQRQRORLEEDIOESRFYRGYGLGHCLQARDGSGMGSIGTSPQAPAPLIR 420
QY 421 EGETTRKLYVD 431
Db 421 EGETTRKLYVD 431
```

RESULT 2  
US-10-075-846-2

```
; Sequence 2, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT E
; FILE REFERENCE: D0079 NP
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 417
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-075-846-2
```

Query Match 96.0%; Score 2163; DB 9; Length 417;  
Best Local Similarity 96.5%; Pred. No. 3.4e-198;  
Matches 416; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

```
QY 1 MTTLPATLSFLMTLTPQVLLRVALAKEEYKSGTGSQPMSPDFLDKLMGRTSGYDA 60
Db 1 MTTLPATLSFLMTLTPQVLLRVALAKEEYKSGTGSQPMSPDFLDKLMGRTSGYDA 60
QY 61 RIRPNFKGPPVAVNTCNIFINSFSYTKTMDYRVAVNFLRQWMDPRLSYREYPPDSDLD 120
Db 61 RIRPNFKGPPVAVNTCNIFINSFSYTKTMDYRVAVNFLRQWMDPRLSYREYPPDSDLD 120
QY 121 PSMDSIMKPDLPFANEKANFHEVTTDNKLLIRKNGNVLYSIRLTILSLCMLDKNP 180
Db 121 PSMDSIMKPDLPFANEKANFHEVTTDNKLLIRKNGNVLYSIRLTILSLCMLDKNP 180
QY 181 MDIOTCTMOLESSLILCSPLSLISVGYTMDLVFEMLEDAPAVOVAGLTLPOFILRD 240
Db 181 MDIOTCTMOLESSLILCSPLSLISVGYTMDLVFEMLEDAPAVOVAGLTLPOFILRD 240
QY 241 EKDLGCTKHYNTGKFTCLEVGFHLEKRGYLLIOMYIPSLIIVLISWSEFIMNDAAPA 300
Db 241 EKDLGCTKHYNTGKFTCLEVGFHLEKRGYLLIOMYIPSLIIVLISWSEFIMNDAAPA 300
QY 301 RVGLGITTIVLMTTQSSGSRASLPVSYKAIIDIMAVCLLFVFAALLEYAAINVSROH 360
Db 301 RVGLGITTIVLMTTQSSGSRASLPVSYKAIIDIMAVCLLFVFAALLEYAAINVSROH 360
QY 361 KEFIRLRQRQRORLEEDIOESRFYRGYGLGHCLQARDGSGMGSIGTSPQAPAPLIR 420
Db 361 KEFIRLRQRQRORLEEDIOESRFYRGYGLGHCLQARDGSGMGSIGTSPQAPAPLIR 420
QY 421 EGETTRKLYVD 431
Db 421 EGETTRKLYVD 431
```

RESULT 3  
US-10-075-846-13

```
; Sequence 13, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; FILE REFERENCE: D0079 NP
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-075-846-13
```

Query Match 74.7%; Score 1682.5; DB 9; Length 452;  
Best Local Similarity 79.0%; Pred. No. 3.1e-152;  
Matches 324; Conservative 29; Mismatches 36; Indels 21; Gaps 5;

```
QY 24 RVALAKE-EVKSQSGTGSQPMSPDFLDKLMGRTSGYDARIRPNFKGPPVAVNTCNIFINSF 82
Db 23 RFAFCQKDHDSRGKQPSQTLSPSDFLDKLMGRTSGYDARIRPNFKGPPVAVNTCNIFINSF 82
QY 83 SSVTKTMDYRVAVNFLRQWMDPRLSYREYPPDSDLDLPMSLDSIMKPDLPFANEKANF 142
Db 83 GSVTETMDYRVAVNFLRQWMDPRLSYREYPPDSDLDLPMSLDSIMKPDLPFANEKANF 142
QY 143 HEVTTDNKLLIRKNGNVLYSIRLTILSLCMLDKNFMNDIOTCTMOLESSLILCSPLS 202
Db 143 HDVTTDNKLLIRKNGKVLISIRLTILSLCSPDLDKNFMDVOTCTMOLE----- 191
QY 203 LSLSVGYTMDLVFEMLEDAPAVOVAGLTLPOFILRDEKDLGCTKHYNTGKFTCLEV 262
Db 192 ---SFGYTMNDLIFEMLDGP-VQVAREGLTLPOFILRKEKEKELGYCTKHYNTGKFTCLEV 247
QY 263 FHLEROMGYLLIOMYIPSLIIVLISWSEFIMNDAAPARVGLGITTIVLMTTQSSGSRAS 322
Db 248 FHLEROMGYLLIOMYIPSLIIVLISWSEFIMNDAAPARVAGLITTVLMTTQSSGSRAS 307
QY 323 LPKVSYKAIIDIMAVCLLFVFAALLEYAAINVSROHKEFIRLRQRQRORLEEDIOE 382
Db 308 LPKVSYKAIIDIMAVCLLFVFAALLEYAAINVSROHKEFIRLRQRQRORLEEDIOE 367
QY 383 SRFYRGYGLGHCLQARDGSGMGSIGTSPQAPAPLIREGETTRKLYVD 431
Db 368 SRFNFSGYGMGHCLQVKGDTAVKATPAMPDLQPP-----KDDGAIKKKKYVD 413
```

RESULT 4  
US-10-075-846-12

```
; Sequence 12, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; FILE REFERENCE: D0079 NP
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 337
; TYPE: PRT
```

ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (322)..  
OTHER INFORMATION: wherein "X" is any amino acid.  
US-10-075-846-12

Query Match 74.5%: Score 1678; DB 9; Length 337;  
Best Local Similarity 92.0%: Pred. No. 5,5e-152;  
Matches 323; Conservative 7; Mismatches 7; Indels 14; Gaps 1;

OY 25 VALAKEVKSSTKSGQSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTNCNIFINSFS  
DB 1 VALAKEVKSGLKSGQSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTNCNIFINSFS  
OY 85 VTKTMDYRVAVFLRQOMNDPRLSYREYPPDSLDLPSMDSIMKPDLEFFANEGANHE 144  
DB 61 VTEETMDYRVAVFLRQOMNDPRLSYREYPPDSLDLPSMDSIMKPDLEFFANEGANHE 120  
OY 145 VTTDNKLLRIFKNGNVLSIRLTLLSCMLDKNFPMDIOTCTMOLESSILCSPLSLS 204  
DB 121 VTTDNKLLRIFKNGNVLSIRLTLLSCMLDKNFPMDIOTCTMOLE----- 167  
OY 205 LSVGYTKMDLVEFELEDAPAVAEGTLTPOFILRDEKDLGCTKHNTGKFTCIEVKFH 264  
DB 168 -SFGYTNMDLVEFELEDAPAVAEGTLTPOFILRDEKDLGCTKHNTGKFTCIEVKFH 226  
OY 265 LEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTTLMTTOSGSRASLP 324  
DB 227 LEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTTLMTTOSGSRASLP 286  
OY 325 KVSIVKAIIDIMAVCLLEFVALLEAYAINFVSROKHEFIRRRORROL 375  
DB 287 KVSIVKAIIDIMAVCLLEFVALLEAYAVNFVSROKHEFIRRRORROL 337

RESULT 5  
US-10-075-846-15  
Sequence 15, Application US/10075846  
Publication No. US20030032608A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT H  
FILE REFERENCE: D0079 NP  
CURRENT APPLICATION NUMBER: US/10/075, 846  
PRIOR FILING DATE: 2002-02-13  
CURRENT APPLICATION NUMBER: US 60/269, 535  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 15  
LENGTH: 312  
TYPE: PRT  
ORGANISM: homo sapiens  
US-10-075-846-15

Query Match 72.7%: Score 1638; DB 9; Length 312;  
Best Local Similarity 100.0%: Pred. No. 3,3e-148;  
Matches 312; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 PSDFLDKLMGRTSGYDARIRPNFKGPPVAVTNCNIFINSFSVTCTMDYRVAVFLRQOMN 103  
DB 1 PSDFLDKLMGRTSGYDARIRPNFKGPPVAVTNCNIFINSFSVTCTMDYRVAVFLRQOMN 60  
OY 104 DPLRISYREYPPDSLDLPSMDSIMKPDLEFFANEGANHEVTTDNKLLRIFKNGNVLS 163  
DB 61 DPLRISYREYPPDSLDLPSMDSIMKPDLEFFANEGANHEVTTDNKLLRIFKNGNVLS 120  
OY 164 IRLTLLISCLMDLKNFMDIOTCTMOLESSILCSPLSLSLVGYTKMDLVEFELEDA 223  
DB 121 IRLTLLISCLMDLKNFMDIOTCTMOLESSILCSPLSLSLVGYTKMDLVEFELEDA 180

OY 224 AVVAEGTLTPOFILRDEKDLGCTKHNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 283  
DB 181 AVVAEGTLTPOFILRDEKDLGCTKHNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 240  
OY 284 VILSWVSFMINMDAPARVGLGITTTLMTTOSGSRASLPKVSIVKAIIDIMAVCLLFV 343  
DB 241 VILSWVSFMINMDAPARVGLGITTTLMTTOSGSRASLPKVSIVKAIIDIMAVCLLFV 300  
OY 344 FALLEAYAINF 355  
DB 301 FALLEAYAINF 312

RESULT 6  
US-10-075-846-11  
Sequence 11, Application US/10075846  
Publication No. US20030032608A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT  
FILE REFERENCE: D0079 NP  
CURRENT APPLICATION NUMBER: US/10/075, 846  
PRIOR FILING DATE: 2002-02-13  
CURRENT APPLICATION NUMBER: US 60/269, 535  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 81  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 11  
LENGTH: 465  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-075-846-11

Query Match 71.2%: Score 1602.5; DB 9; Length 465;  
Best Local Similarity 73.8%: Pred. No. 1,4e-144;  
Matches 312; Conservative 35; Mismatches 43; Indels 33; Gaps 6;

OY 22 LLRVALAKEEVKSGTSGQSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTNCNIFINS 81  
DB 21 LLSLVAKKETDSARSAPSPSPDFLDKLMGRTSGYDARIRPNFKGPPVAVTNCNIFINS 80  
OY 82 FSSVYTKTMDYRVAVFLRQOMNDPRLSYREYPPDSLDLPSMDSIMKPDLEFFANEGAN 141  
DB 81 FGSIAETMDYRVAVFLRQOMNDPRLSYREYPPDSLDLPSMDSIMKPDLEFFANEGAN 140  
OY 142 FHEVTTDNKLLRIFKNGNVLSIRLTLLSCMLDKNFPMDIOTCTMOLESSILCSPL 201  
DB 141 FHEVTTDNKLLRIFKNGNVLSIRLTLLSCMLDKNFPMDIOTCTMOLE----- 190  
OY 202 SLISVGYTKMDLVEFELEDAPAVAEGTLTPOFILRDEKDLGCTKHNTGKFTCIEV 261  
DB 191 -SFGYTNMDLVEFELEDAPAVAEGTLTPOFILRDEKDLGCTKHNTGKFTCIEV 245  
OY 262 KHLEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTTLMTTOSGSRAS 321  
DB 246 RHLEROMGYLLIOMYIPSLIIVLSWVSFMINMDAPARVGLGITTTLMTTOSGSRAS 305  
OY 322 SLRVSIVKAIIDIMAVCLLEFVALLEAYAINFVSROKHEFIRRRORROL----- 376  
DB 306 SLRVSIVKAIIDIMAVCLLEFVALLEAYAVNFVSROKHEFIRRRORROL----- 365  
OY 377 -----EDIOESRFYRGYGLGICLOARDGPMGSGIYSPQPPAPLL-REGFTTKL 428  
DB 366 FYRFSMDDEVRESFSTAYAGMGPCLQAKGOMPKG-----PNHPVQVMKSPDEMKV 420  
OY 429 YVD 431  
DB 421 FID 423

RESULT 7  
US-10-075-846-10

; Sequence 10, Application US/10075846  
; Publication No. US20030032608A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT E  
; FILE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 10  
; LENGTH: 449  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-075-846-10

Query Match 70.2%; Score 1580.5; DB 9; Length 449;  
Best Local Similarity 73.4%; Pred. No. 1.7e-142;

Matches 307; Conservative 38; Mismatches 50; Indels 23; Gaps 7;

QY 17 LPQVLLKVALAKEVSGTSGSPMSDFLDKLMGRTSGVDARIRNFKGPPVNVTCN 76  
DB 11 LSGAIVFSLASKAEARSAFTKPMSPSDFLDKLMGRTSGVDARIRNFKGPPVNVSCN 70  
QY 77 IFINSSSVTKTMDYRVNVLRLQOMNDPRLSREYPPDSDLDPSMDSIMKPOLPFRAN 136  
DB 71 IFINSGSIATETMYRNIFRLROOMNDPRLANRYPPDSDLDPSMDSIMKPOLPFRAN 130  
QY 137 EKGANFHEVTDNKLIRFKNGNVLSIRLLILSLMDLKNFPMDIOTCTMOLESSITL 196  
DB 131 EKGAFHEITDNKLIRSRNGNVLSIRITLTCLACPMDLKNFPMDOVOTCIQLE----- 185  
QY 197 CSPPLSLSLVGYTKMDLVEFMLEDAPOVAEGTLTPOFILRDKDLGCCCKHNTGKF 256  
DB 186 -----SFGYTMNDLFLFEMOEG-AYOVADGLTLPOFILKEKDLRYCKRHYNTGKF 235  
QY 257 TCIEVKFHLEROMGYLLIOMYIPSLIYLIVLSWSEFWINMDAPARVGLITVLTMTQS 316  
DB 236 TCIEARFHLEROMGYLLIOMYIPSLIYLIVLSWISFWINMDAPARVGLITVLTMTQS 295  
QY 317 SGRASLKRVSIVKADIDIMAVCLFVFAALLEVAIAINFEVSQKHETRLRRORORLE 376  
DB 296 SGRASLKRVSIVKADIDIMAVCLFVFAALLEVAIAINFEVSQKHETRLRRORORLE 353  
QY 377 EDIIESRFYFPGYGLG-HCLQARDGPMEG---SGIYSPQPPARLLHEGETTRKLY 430  
DB 354 EDEAGEGRNFSAYMGAPACIDAKDISYKGNANSNTNP-PPAP-SKSPPEMRKLF 409

RESULT 8  
US-10-075-846-14

; Sequence 14, Application US/10075846  
; Publication No. US20030032608A1

; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT E  
; FILE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF  
; FILE REFERENCE: D0079 NP  
; CURRENT APPLICATION NUMBER: US/10/075,846  
; CURRENT FILING DATE: 2002-02-13  
; PRIOR APPLICATION NUMBER: US 60/269,535  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 14  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-075-846-14

Query Match 68.8%; Score 1549; DB 9; Length 298;  
Best Local Similarity 95.2%; Pred. No. 9.7e-140;  
Matches 297; Conservative 0; Mismatches 1; Indels 14; Gaps 1;

QY 44 PSDFLDKLMGRTSGVDARIRNFKGPPVNVTCNIFINSSSVTKTMDYRVNVLROOM 103  
DB 1 PSDFLDKLMGRTSGVDARIRNFKGPPVNVTCNIFINSSSVTKTMDYRVNVLROOM 60  
QY 104 DPLRSREYPPDSDLDPSMDSIMKPOLPFRANEGANFHEVTDNKLIRFKNGNVLS 163  
DB 61 DPLRSREYPPDSDLDPSMDSIMKPOLPFRANEGANFHEVTDNKLIRFKNGNVLS 120  
QY 164 IRLTLILSLMDLKNFPMDIOTCTMOLESSITLSPPLSLSLVGYTKMDLVEFMLEDAP 223  
DB 121 IRLTLILSLMDLKNFPMDIOTCTMOLE-----SFGYTKMDLVEFMLEDAP 166  
QY 224 AVQVAEGTLTPOFILRDEKDLGCCCKHNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 283  
DB 167 AVQVAEGTLTPOFILRDEKDLGCCCKHNTGKFTCIEVKFHLEROMGYLLIOMYIPSLI 226  
QY 284 VILSVSEFWINMDAPARVGLITVLTMTQSSGRASLKRVSIVKADIDIMAVCLFV 343  
DB 227 VILSVSEFWINMDAPARVGLITVLTMTQSSGRASLKRVSIVKADIDIMAVCLFV 286  
QY 344 FAALLEVAIANF 355  
DB 287 FAALLEVAIANF 298

RESULT 9  
US-09-969-844-14

; Sequence 14, Application US/09969844  
; Publication No. US20020192776A1  
; GENERAL INFORMATION:

; APPLICANT: XIAO-ZHOU Michelle Wang  
; APPLICANT: Xavier Georges Sarda

; APPLICANT: Michael David Tomalski  
; APPLICANT: Vincent Paul Mary Wingate

; TITLE OF INVENTION: Helicobacter Glutamate Receptor  
; FILE REFERENCE: A32815-1 072667, 0178  
; CURRENT APPLICATION NUMBER: US/09/969,844  
; CURRENT FILING DATE: 2001-10-03  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 450  
; TYPE: PRT  
; ORGANISM: Helicobacter virescens  
US-09-969-844-14

Query Match 32.7%; Score 737; DB 9; Length 450;  
Best Local Similarity 45.7%; Pred. No. 6.7e-62;  
Matches 158; Conservative 53; Mismatches 111; Indels 24; Gaps 7;

QY 31 EYKSGTSGSPMSDFLDKLMGRTSGVDARIRP---NFKGPPVNVTCNIFINSSSVTK 87  
DB 1 EYKSGTSGSPMSDFLDKLMGRTSGVDARIRP---NFKGPPVNVTCNIFINSSSVTK 87  
QY 26 EGMNGKINFREREKQIIDLIG-EGRYDARIRPSGINGTGPVAVSVINIFRSISKID 84  
DB 86 TTMDBRVAVFLRQOMNDPRLSREYPPDSDLDPSMDSIMKPOLPFRANEGANFHEVTT 147  
QY 85 VTMEISVOLTFREQWDLERLKNNGIRKLYTLTEARVWMPDLFESNEKEGHHNITM 144  
DB 148 DNKLIRFKNGNVLSIRLLILSLMDLKNFPMDIOTCTMOLESSITLSPPLSLSLV 207  
QY 145 PNVIYIRIFPNNGVLSIRISLTLSCPMNLKLYPLDKQCSLRM-----ASY 190  
DB 191 GTTDDLVFLMKEGDP-VQVVKNLILPFTLT--EKFLDYCNKNTNGEYCLKDLDFK 247  
QY 208 GYTKMDLVEFMLEDAPOVAEGTLTPOFILRDEKDL-GCCTKHNTGKFTCIEVKFHL 266  
DB 191 GTTDDLVFLMKEGDP-VQVVKNLILPFTLT--EKFLDYCNKNTNGEYCLKDLDFK 247  
QY 267 RQMGYYLLIOMYIPSLIYLIVLSWSEFWINMDAPARVGLITVLTMTQSSGRASLKR 326  
DB 248 REFSTYLLIYIPCCMLIVSVSWFMDQGAIVARVLLGVTTLTMTATQSSGINASLPV 307



```

1      REGISTRATION NUMBER: 45,243
2      REFERENCE/DOCKET NUMBER: T1292
3      TELECOMMUNICATION INFORMATION:
4      TELEPHONE: 732-594-1307
5      TELEFAX: 732-594-4720
6      TELEX: UNKNOWN
7      INFORMATION FOR SEQ ID NO: 12:

```

```

QY      48 LDKLMGRTSGVDARLRP-NFKGPVNVVTCNFIINSSSVTKTMDYRVNVPFRQQWMDPR 106
      || :: || || | : | | | : : || : || | : :: || | | |
Db      53 LDNIILV---SYDRRIPTGGHNLNVPVYVKCEIYLRKSGAVNPATMDVDVLYRQTWTDLR 109

```

Job time : 37.6756 secs

```
QY 107 LSYREXPDDSLDLDPSMDSIMKPDLEFANEKGFHEVTTDNKLLRIFKNGNVLYSIR 165
      :      :      :      :      :      :      :      :      :      :
Db 110 MKNANL-TRSLDNDPRLKWKVPDYFPNAKHGEQFVTPNVLRLIRYPTGDILYMLR 168
QY 166 LTIILSCMDLKNFPMDIOTCTMOLESSILCSPLSLSVGYTKMDLVEWLEDAPAV 225
      :      :      :      :      :      :      :      :      :      :
Db 169 LKLFSCMMHMERYPIDRQVCSIELASF-----KTEVELOW-GNAEAV 213
QY 226 QVAGEGLTPQFIRDEKDGCCCKHNTGCKFCIEVKFHLEROMGYLLIOMYIPSLIYI 285
      :      :      :      :      :      :      :      :      :      :
Db 214 TMSGLKMAQFELQ-QSLTKCSGAFQIGEYSCLRALNLRKSRIGHLVOSYLPSTLIYV 272
QY 286 LSNVSEFINMDAPARYGIGITVTLTWTQSSGRASLPKVSYKAIIDIMAVCLLFVFA 345
      :      :      :      :      :      :      :      :      :      :
Db 273 VSNVSEFLDVAIPARTIGVTLTLTISSESDHQANLAPSYKALDVWMTGCTMVFFA 332
QY 346 ALLEYAIAINFVSRO 359
      :      :      :      :      :      :      :      :      :      :
Db 333 AVLETFVSYLARR 346
```

## RESULT 15

```
US-09-510-662A-39
; Sequence 39, Application US/09510662A
; Patent No. US20020155125A1
; GENERAL INFORMATION:
; APPLICANT: Dillion, Davin C.
; APPLICANT: Wang, Craig H.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.491C1
; CURRENT APPLICATION NUMBER: US/09/510.662A
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 39
; LENGTH: 440
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-510-662A-39
```

Query Match 29.0%; Score 652.5; DB 9; Length 440;

Best Local Similarity 39.1%; Pred. No. 7.6e-54;

Matches 133; Conservative 70; Mismatches 110; Indels 27; Gaps 7;

```
QY 55 TSGYDARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRVNVLROQWMDPRLSYREYPD 114
      :      :      :      :      :      :      :      :      :      :
Db 45 TAGYNKFLRPNFGEGEPQIALTLDIASISSISESNMDYTATYLRQWMDQRLVFE--GN 102
QY 115 DSDLDPSMDSIMKPDLEFANEKGFHEVTTDNKLLRIFKNGNVLYSIRLTIILSCLM 174
      :      :      :      :      :      :      :      :      :      :
Db 103 KSFLLDLRLVEFLWNPDTYVESKSKFLHEVTGNRLIRLESNGTVLYALRLITTVACNM 162
QY 175 DLKNFMDIOTCTMOLESSILCSPLSLSVGYTKMDLVEWLEDAPAVOVAEGTLTP 234
      :      :      :      :      :      :      :      :      :      :
Db 163 DLSKYPMDTQCKQLE-----SMGCDNDGVEFTWLRGNDVSYGLHLRLA 208
QY 235 QFIRDEKDLGCGCK-HYNTGKFTCIEVKFHLEROMGYLLIOMYIPSLIYIISVSWFSWI 293
      :      :      :      :      :      :      :      :      :      :
Db 209 QYTI--ERYFTLYTRSOQETGNTRYLRLOFELRNVLYFLETVVPSFPLVLSWVSFWI 266
QY 294 NMDAAPARVIGITVTLTWTQSSGRASLPKVS-YVKATIDIMAVCLLFVFAALLEYAA 352
      :      :      :      :      :      :      :      :      :      :
Db 267 SLDSPVPAKTCIGVTVLMTTLMIGSRTSLPNTNCFKALDVIYIGICFSFVGFALLEYAV 326
QY 353 INFVSRQHKKEFIRLRQRQRLE---DIQESREYFR 388
      :      :      :      :      :      :      :      :      :      :
Db 327 AHYSLSQ---MAKDRGTTRKEVEVSITNINSSISPK 363
```

Search completed: June 25, 2003, 17:18:50

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:14:25 ; Search time 23.1409 Seconds  
(without alignments)  
548.002 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPAFLSLFLMTLPGQ.....PQPPAPLLREGETTRKLYVD 431

Scoring table:  
BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/plodata/1/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/plodata/1/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/plodata/1/1aa/6A\_COMB.pep: \*  
4: /cgn2\_6/plodata/1/1aa/6B\_COMB.pep: \*  
5: /cgn2\_6/plodata/1/1aa/PCUTUS\_COMB.pep: \*  
6: /cgn2\_6/plodata/1/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	745	33.1	617	1	US-08-137-614A-24 Sequence 24, Appl
2	737	32.7	450	4	US-09-592-891A-14 Sequence 14, Appl
3	731.5	32.5	455	4	US-09-130-339-2 Sequence 2, Appl
4	730.5	32.4	456	1	US-08-435-933-6 Sequence 6, Appl
5	730.5	32.4	456	5	PCT-US96-06035-6 Sequence 6, Appl
6	730	32.4	423	4	US-09-627-650B-19 Sequence 19, Appl
7	726	32.2	637	3	US-08-072-064-8 Sequence 8, Appl
8	724.5	32.2	459	4	US-09-002-361-6 Sequence 8, Appl
9	724.5	32.2	467	4	US-09-002-361-3 Sequence 6, Appl
10	724.5	32.2	481	4	US-09-002-361-5 Sequence 3, Appl
11	724.5	32.2	496	4	US-09-002-361-2 Sequence 5, Appl
12	720	32.0	488	1	US-08-554-659-2 Sequence 2, Appl
13	720	32.0	488	1	US-08-554-659-4 Sequence 4, Appl
14	720	32.0	637	3	US-08-072-064-1 Sequence 1, Appl
15	720	32.0	637	3	US-08-072-064-4 Sequence 1, Appl
16	720	32.0	637	3	PCT-US92-08558-1 Sequence 1, Appl
17	718	31.9	637	5	US-08-072-064-6 Sequence 1, Appl
18	717	31.8	496	1	US-08-137-614A-2 Sequence 2, Appl
19	717	31.8	496	2	US-08-768-301-2 Sequence 2, Appl
20	714.5	31.7	383	4	US-09-627-650B-20 Sequence 20, Appl
21	712	31.6	474	1	US-08-417-330A-20 Sequence 20, Appl
22	705	31.3	617	1	US-08-137-614A-26 Sequence 26, Appl
23	691.5	30.7	474	1	US-08-459-100A-4 Sequence 4, Appl
24	683.5	30.4	487	4	US-09-627-650B-13 Sequence 13, Appl
25	680.5	30.4	487	4	US-09-436-063C-13 Sequence 13, Appl
26	680.5	30.2	617	1	US-08-137-614A-25 Sequence 25, Appl
27	678.5	30.1	452	4	US-08-809-802-8 Sequence 12, Appl

28	673	29.9	487	1	US-08-249-112-4 Sequence 4, Appl
29	673	29.9	487	5	PCT-US95-06556-4 Sequence 4, Appl
30	670.5	29.8	510	1	US-08-249-112-3 Sequence 3, Appl
31	670.5	29.8	510	5	PCT-US95-06556-3 Sequence 3, Appl
32	663	29.4	506	4	US-09-627-650B-15 Sequence 15, Appl
33	663	29.4	506	4	US-09-436-063C-15 Sequence 15, Appl
34	657	29.2	487	4	US-09-627-650B-14 Sequence 14, Appl
35	657	29.2	487	4	US-09-436-063C-14 Sequence 14, Appl
36	656	29.1	475	4	US-09-627-650B-16 Sequence 16, Appl
37	656	29.1	475	4	US-09-436-063C-16 Sequence 16, Appl
38	652.5	29.0	440	1	US-08-459-100A-2 Sequence 2, Appl
39	652.5	29.0	440	5	PCT-US94-09589-2 Sequence 2, Appl
40	644.5	28.6	453	1	US-08-417-330A-18 Sequence 18, Appl
41	628	27.9	686	2	US-08-768-301-4 Sequence 4, Appl
42	626	27.8	467	1	US-08-459-100A-3 Sequence 3, Appl
43	619.5	27.5	554	4	US-08-809-802-8 Sequence 8, Appl
44	618	27.4	451	1	US-08-417-330A-12 Sequence 12, Appl
45	610.5	27.1	492	1	US-08-417-330A-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1  
US-08-137-614A-24  
Sequence 24, Application US/08137614A  
Patent No. 5487676  
GENERAL INFORMATION:  
APPLICANT: Soderlund, David M.  
APPLICANT: Knipfle, Douglas C.  
APPLICANT: Henderson, Joseph E.  
TITLE OF INVENTION: Gene Encoding An Insect  
TITLE OF INVENTION: Gamma-Aminobutyric Acid (GABA) Receptor Subunit  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/137,614A  
FILING DATE: 15-OCT-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Timilan, Susan J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 19603/120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716)263-1636  
TELEFAX: (716)263-1600  
INFORMATION FOR SEQ ID NO. 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 617 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-137-614A-24

Query Match 33.1%, Score 745, DB 1, Length 617;  
Best Local Similarity 42.2%, Pred. No. 2.9e-70;  
Matches 146, Conservative 71, Mismatches 95, Indels 34, Gaps 6;  
QY 48 LDKMGGRSGVDARIRPFKGPVAVTCNIFNSSTYTKTTMDRVAVFLRQGNDRRL 107  
DB 63 IDSLT--GGYDIRLRPSFGAPLEIGIEVLASPDISEVMDVYTIMYINQWRDERL 119

QY 108 SYREYDDSLDLP-----SMDSIMKDPLEFANEKANFHEVTTDNKLRIF 155  
Db 120 QFXIF-NESLTDGENRSVXTXMTLTGAFAEKIMWDEFLANDKNSFLHIDFEKKNMVRLY 178  
QY 156 KNGNLYSIRLTLISCLMDLKNFMDIQTCTMOLESSILCSPISLSLSTGYTKMDLY 215  
Db 179 GNGSLVYGMRTTTLACMKMDLHNNPLDHQECVEIE-----SYGYTMDIY 224  
QY 216 FEMLEDAVAOVAEGTLPOFILRDEKDCCTKHYNKGFCEVFKFHLEROMGYLIDQ 275  
Db 225 LYMLNDRGAVYGEVSVLPOFSITNYKATIN-KIEELSTGDIORLSLITOLQKNIGFTIFQ 283  
QY 276 MYPSLLIVILSVSWSEFMINMDAPARVGLITVTLMTTQSSGSRASLPKVSYKAIDIM 335  
Db 284 TYPSPILLVIMLSWSFMINHEATSARVALGITVTLMTTISNGVRSPLRISYKAIDY 343  
QY 336 MAVCLLFVPALEVAIAINFV---SROKKEFIRLRORRQRL 378  
Db 344 LYMCVEFVFAALLEYAANVYTWGARAKRKAKRLERATSVKRVD 389

## RESULT 2

US-09-592-891A-14  
Sequence 14, Application US/09592891A  
Patient No. 6329174  
GENERAL INFORMATION:  
APPLICANT: Xiao-Zhou Michelle Wang  
APPLICANT: Xavier Georges Sarda  
APPLICANT: Michael David Tomalski  
APPLICANT: Vincent Paul Mary Wingate  
TITLE OF INVENTION: Heliothis glutamate Receptor  
FILE REFERENCE: A32815 072667.0118  
CURRENT FILING DATE: 2001-01-22  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 450  
TYPE: PRT  
ORGANISM: Heliothis virescens  
US-09-592-891A-14

Query Match 32.7%; Score 737; DB 4; Length 450;  
Best Local Similarity 45.7%; Pred. No. 1.3e-69;  
Matches 158; Conservative 53; Mismatches 111; Indels 24; Gaps 7;

QY 31 EVKSGTKSOPKSPSPDFDKLMGRTSGYDARIRP---NFKGPPVAVTCNIFNSPSTYK 87  
Db 26 EOMNGCKINFRKEKEQIIDDQILG-PGRYDARIRPSCINGTDCPAAVSVNIIVRSISKIDD 84  
QY 88 TTMDFRVAVFLRQONNDRLSYREYDDSLDLPMSLDSIMKDPLEFANEKANFHEVTT 147  
Db 85 VTMEYSVOLTFREQWLDERLKFNNLGGRLKYLTLEARNVWMPDLFFSNEKGRPHNIIM 144  
QY 148 DNKLIRFKNGVLYSIRLTLISCLMDLKNFMDIQTCTMOLESSILCSPISLSLSTGY 207  
Db 145 PNVYIRIRPNNGVLYSIRLTLISCLMDLKNFMDIQTCTMOLESSILCSPISLSLSTGY 190  
QY 208 GYTMDFVEMLEDAVAOVAEGTLPOFILRDEKDC-GCCCKHYNTGKFTCEVFKFHE 266  
Db 191 GATTDLDVFLMKEGDP-VQVYKNLHLPRTL--EKLTDYCKSKTNGTGYSLKAYDLLEPK 247  
QY 267 ROMGYLLQNTYIPSLIIVLSVSWSEFMINMDAPARVGLITVTLMTTQSSGSRASLPK 326  
Db 248 REFSYLLIOIYIPCCMLVIVSWSEFMDOGANPARVLLGVTTLLTMAOTSSINSLSPV 307  
QY 327 SYVAIDIMAVCLLFVPALEVAIAINFVSR--OHKEFIRLRORRQ 370  
Db 308 SYTKAIDVWGTGCLTFVFGALLESRFVNVASRSDMHRENMKKARRE 353

RESULT 3  
US-09-130-339-2

Sequence 2, Application US/09130339  
Patient No. 6358701  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Paresse, Philip S.  
APPLICANT: Warmke, Jeffrey W.  
APPLICANT: Eter, Adrian  
APPLICANT: Cohen, Charles J.  
APPLICANT: Brochu, Richard M.  
TITLE OF INVENTION: DNA MOLECULES ENCODING CLENOCEPHALIDES  
FILE REFERENCE: 20029  
CURRENT FILING DATE: 1998-08-06  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 455  
TYPE: PRT  
ORGANISM: clenocephalides felis  
US-09-130-339-2

Query Match 32.5%; Score 731.5; DB 4; Length 455;  
Best Local Similarity 43.3%; Pred. No. 4.9e-69;  
Matches 157; Conservative 60; Mismatches 111; Indels 35; Gaps 9;

QY 24 RVALKEEVKSGTKSOPKSPSPDFDKLMGRTSGYDARIRP---FKGPPVAVTCNIF 78  
Db 27 RUGGKKEKFRAKER-----QVLDQILG-PGHTDARIRPSCINGTDCGPTV-VAAVITY 76  
QY 79 INSFSVTKTMDRVAVFLRQONNDRLSYREYDDSLDLPMSLDSIMKDPLEFANEK 138  
Db 77 LRSIEIDYKMEYSVOLTFREQWDERLKFNDGGRKLYLTLEARNVWMPDLFFSNEK 136  
QY 139 GANFHEVTTDNKLIIFKNGVLYSIRLTLISCLMDLKNFMDIQTCTMOLESSILCS 198  
Db 137 EGFHNIIMPVYIRIFPYGSVLYSIRLTLISCLMDLKNFMDIQTCTMOLESSILCS 188  
QY 199 PLPSLISVGYTMKLVFEMLEDAVAOVAEGTLPOFILRDEKDC-GCCCKHYNTGKFT 257  
Db 189 -----ASYGWTNDLVFLMKEGDP-VQVYKNLHLPRTL--EKLTDYCKSKTNGTGY 239  
QY 258 CIEVKFHLEROMGYLLQNTYIPSLIIVLSVSWSEFMINMDAPARVGLITVTLMTTQSS 317  
Db 240 CLKVLDLFRKERSYLLIOIYIPCCMLVIVSWSEFMDOGANPARVLLGVTTLLTMAOT 299  
QY 318 GSRASLPVSYKADIMAVCLLFVPALEVAIAINFVSR--OHKEFIRLRORRQRL 375  
Db 300 GINASLPVSYKADIMAVTGYCLTFVFGALLEFALVNTASRSDMHRENMKKRRLEBOA 359  
QY 376 EED 378  
Db 360 SLD 362

## RESULT 4

US-08-435-933-6  
Sequence 6, Application US/08435933  
Patient No. 5693492  
GENERAL INFORMATION:  
APPLICANT: Cully, Doris F.  
APPLICANT: Arena, Joseph P.  
APPLICANT: Paresse, Philip S.  
APPLICANT: Liu, Ken K.  
TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE CHANNELS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John W. Wallen III  
STREET: 126 East Lincoln Avenue  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: US



US-08-072-064-8  
 US-08-072-064-8  
 Sequence 8, Application US/08072064  
 Patent No. 6008046  
 GENERAL INFORMATION:  
 APPLICANT: FRENCH-CONSTANT, RICHARD H.  
 APPLICANT: JACKSON, MEYER B.  
 TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING  
 NUMBER OF SEQUENCES: 20  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: PETER G. CARROLL  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/072,064  
 FILING DATE: 19930602

```

OY 5 VPPTLSFLLMTLPGCVLLRLVALAKE-----VKSQTKSGQMSPSDLKLMGRTS 56
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 46 LPRFT - PLUTLW-----LAINMALLIQEYGHKRRIHTYQATGGGSMLGPNVLSALDSRV 99
OY 57 GYDARLRPNKEGPVNVTCNIIINFSSTVKTMTDYRYNVNELRQOANDPLRSTYEDP-D 115
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 100 SYDKRRVRPNVNGPEVPGVTMYVLSISSSEVLMQFTLDFEYRQFWPTPRLAYRGRVE 155
OY 116 SLDLDSMLDSIKKPDLFEANKGAFNHFVPTDNLKLRFPKGNVLYSIRLLTLLSCMD 175
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 TLSVGSEFLKNITWPTDFEYVNEKOSTFPHATTNBEFIVHHSGSTTNSIRLLTTRASCMP 215
OY 176 LKNFPMQDIOYTCMQLLESSIILCSPLSLSVGYTKMDLVFEMLEDAFPAVAGLTLPQ 235
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 LQYFPMQDQLCHTEIE-----SGGYTMRIQRFMRDGLSSVGMSSVELPQ 265
OY 236 FILRDEKDLGCCKKNHNTKFCIEYKFLHEKQMYLYLQMTIPSLIYLISWYFWINM 295
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 FRYLGHRO - RATEINLTGTNYSRLACEIOFVSMQYVLIQIYIPGLSLTVSWVSFWLNR 325
OY 296 DAAPARVNGIGITVLTMTQSSGSRASLPKVSRYVAIDIMAVCLLFPYFALLLEAFAINF 355
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 325 NAFPARVNGIGITVLTMTLMTLSSTNAALPKISYVSDIVYIGTCYVMVFAALLEAYATGY 368
OY 356 VSNQHKETFLRRRORRRLLEEDITQESHFYRGYGLCHC---LQARDGPEMGSSGITYP 412
  | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 385 MAK-----RIOMRKORFMAIOKIEBOKKOOLDGANOQOANPNPNANVAGC-GGVGVGP 437
OY 413 QPPAP 417
  |
Db 438 GPGGP 442

```

RESULT 8  
US-09-002-361-6  
Sequence 6, Application US/09002361  
Patent No. 6328516  
GENERAL INFORMATION:  
APPLICANT: Halling, Blak  
TITLE OF INVENTION: Lepidopteran GABA-gated Chloroide  
TITLE OF INVENTION: Channels  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210  
CITY: Lawrenceville  
STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002.361  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 459 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-002-361-6

Query Match 32.2%; Score 724.5; DB 4; Length 459;  
Best Local Similarity 37.8%; Pred. No. 2.7e-68;

Matches 153; Conservative 79; Mismatches 120; Indels 53; Gaps 8;

58 YDARIRPNFKGPVNVTCNIFNSFSVTKTMDYRVNVLRLQGMNDRLSYREYD-DS 116  
117 LDLDPSMDSIWKPDLEFANEKGFHEVTTDNKLLRIFKNGVLSIRLTLISCLMDL 176  
85 LSGSEFIKNIWPDLEFVNEKOSYFHIAATSNFIRIHSGSITRSIRLITFASCPMNL 144  
177 KNFPMIDOTCTMOLESSIICSPPLSLSVGYTMKDLFEMLDAPVQVABSLTLPQF 236  
145 QYPMIDQDLCHIELE-----SFGYTMRIKRWNGSPVSSVSSLPQF 190  
237 ILRDEKDLG---CCTKHYNTGKFTCIEVKFHLEROMGYLLIOMYISLILVLSWSPW 292  
191 ----KVLGHRQAMEISLTGTGNVSRILACEIOFVRSMGYLLIOTIISGLIIVISWSPW 245  
293 INMDAPARVGLGTTVLTMTTOSGSRASLPKVSYVKAIDIMNVAVCLLFVFAALLEYAA 352  
246 LNNNAFPAVALGVTYVLTMTTLMSSSTNAALPKISYKSIDVYLGTCFVWVFASLLEYAT 305  
353 INFEVSR---OHKEFIRLRORRORLEEDI--IQESRFYFRGGLGHCLQAR----- 399  
306 VGTMAKRIOMKQKFEVALQIASEKKIPVDCPPVGDPTLSKMGTLRLCPGPRSEVFRK 365  
400 -----DGGPME-----GSGIYSPQPPAPVLRREGETTRKL 428  
366 VHPDKAHSKGTLENTINGRSGAEENPGPPHILRPGKDISKL 410

RESULT 9  
US-09-002-361-3  
Sequence 3, Application US/09002361  
Patent No. 6329516

GENERAL INFORMATION:  
APPLICANT: Halling, Blak  
TITLE OF INVENTION: Lepidopteran GABA-Gated Chloride  
Channels  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dechert Price & Rhoads  
STREET: 997 Lenox Drive, Building 3, Suite 210  
CITY: Lawrenceville

STATE: NJ  
COUNTRY: USA  
ZIP: 08543  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002.361  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Bloom, Allen  
REGISTRATION NUMBER: 29,135  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 609-520-3214  
TELEFAX: 609-520-3259  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 467 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-002-361-3

Query Match 32.2%; Score 724.5; DB 4; Length 467;  
Best Local Similarity 39.6%; Pred. No. 2.8e-68;

Matches 152; Conservative 72; Mismatches 109; Indels 51; Gaps 8;

45 SDFLDKIMGRTSGYDARIRPNFKGPVNVTCNIFNSFSVTKTMDYRVNVLRLQGMND 104  
15 SAIDSL---SVSYDKVRPNYGGPVQVGYTMVLSISSIEVKMDFLDFYFRQFWTD 71  
105 PRLSYREYD-DSLDLDPSMDSIWKPDLEFANEKGFHEVTTDNKLLRIFKNGVLS 163  
72 PRLAYKRTGVEYELTSVSEFIRNIWPDLEFVNEKOSYFHIAATSNFIRIHSGSITRS 131  
164 IRLTLISCLMDLKNFPMIDOTCTMOLESSIICSPPLSLSVGYTMKDLFEMLDAP 223  
132 IRLTITASCMDLOYPFMDROLCHIELE-----SFGYTMRIKRWNGSPN 177  
224 AVOVAEGLTLPQFLRDEKDLG---CCTKHYNTGKFTCIEVKFHLEROMGYLLIOMYIP 279  
178 SVGVSSVSLPQF-----KVLGHRQAMEISLTGTGNVSRILACEIOFVRSMGYLLIOTIIP 232  
280 SLIIVLISWSPWNIINMDAPARVGLGTTVLTMTTOSGSRASLPKVSYVKAIDIMNVAV 339  
233 SGLIIVISWSPWNIINMDAPARVGLGTTVLTMTTLMSSSTNAALPKISYKSIDVYLGTC 292  
340 LLEFVFAALLEYAAINFEVSRQHKFIRLRORRORLEEDIQESRFYFRGGLGHCLQAR 399  
293 FVWVFASLLEYATGYNAK-----RIOMKRORRTAVQKMAEKKMOI----- 334  
400 DGGPMEGSGIYSP-----QPPAP 417  
335 DGGPMSAEPPIPPPTSTLSRPPPP 358

RESULT 10  
US-09-002-361-5  
Sequence 5, Application US/09002361  
Patent No. 6329516

GENERAL INFORMATION:  
APPLICANT: Halling, Blak  
TITLE OF INVENTION: Lepidopteran GABA-Gated Chloride  
Channels  
NUMBER OF SEQUENCES: 43

```

CORRESPONDENCE ADDRESS:
ADDRESS: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002.361
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 481 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-002-361-5

Query Match      32.2%; Score 724.5; DB 4; Length 481;
Best Local Similarity 37.8%; Pred. No. 3e-68;
Matches 153; Conservative 79; Mismatches 120; Indels 53; Gaps 8;

58 YDARIKPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVEFLQONNDPLSYREYPD-DS 116
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
47 YDKRIRPNKGPVPEVGYTMVLSISVSEVLMDFTLDFYRFQFTDPLATKKTGVEY 106
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
117 LDLDPSMLDSIMKPDLEFANEGKGFHEVTTDNKLRIFKNGNVLSIRLTLILSLMDL 176
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
107 LSVGSEFKNIMVPTFFVNEKQSYFHATISNEFRIHYSSTIRSLRITTAQCPMNL 166
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
177 KNEPDIQTCTMOLESSSILCSPLPSLSISVGYTKDLVEFLQONNDPLSYREYPD-DS 236
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
167 QYFPMRDLCHIEI-----SFGYTRDIRYKNNEGPNISVGSSEVSLDQF 212
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
227 ILRDEKDLG---CCTKHYNTEGKFTICIEVKFHELEOMGYLLQNTIPSLIYLISWVSW 292
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
213 -----KVLGHORAMEISLTGNYNSRLACEIOFVSMGYLLQIYIPSLIYLISWVSW 267
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
293 INMDAPARVAGITVTLMTTQSSGSRASLPKVSVAIDIMAVCLLEFVAALLEVAA 352
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
268 LNRNATPARVALGVTTLMTTLMSTNAALPKISYVSIDVYLGTCFMYFASLLEVAT 327
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
333 INEVR-----QKKEITRLRRRQRLQLEEDT-IOESRPFYRGYGLGHCLQAR----- 399
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
328 VGYMAKRIOMKQRFVAIOKIASKEKIPVDCPPVGDPTLTKMGTLGRCPGRPSPSEVRPK 387
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
400 -----DGGPME-----GSGIYSPQAPPLRGEETIRKL 428
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
368 VHDPAHSGGTLENTINGRSGAEENDEGPPHILHPEKDISKL 432
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

Patent No. 6329516
GENERAL INFORMATION:
APPLICANT: Halling, Blaik
TITLE OF INVENTION: Lepidopteran GABA-gated Chloride
NUMBER OF INVENTION: Channels
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESS: Dechert Price & Rhoads
STREET: 997 Lenox Drive, Building 3, Suite 210
CITY: Lawrenceville
STATE: NJ
COUNTRY: USA
ZIP: 08543
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/002.361
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bloom, Allen
REGISTRATION NUMBER: 29,135
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-520-3214
TELEFAX: 609-520-3259
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-002-361-2

Query Match      32.2%; Score 724.5; DB 4; Length 496;
Best Local Similarity 39.6%; Pred. No. 3.1e-68;
Matches 152; Conservative 72; Mismatches 109; Indels 51; Gaps 8;

45 SDFDKLMKRTSGYDARIKPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVEFLQONND 104
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
44 SALDLSL---SVYDKRVRPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVEFLQONND 100
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
105 PLSYREYPD-LSLDPSMLDSIMKPDLEFANEGKGFHEVTTDNKLRIFKNGNVLS 163
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
101 PLATKKTGVEYTLVSGSEFKNIMVPTFFVNEKQSYFHATISNEFRIHYSSTIRSL 160
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
164 ILTLILSLMDLKNFMDIQTCTMOLESSSILCSPLPSLSISVGYTKDLVEFLQONND 223
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
161 IRLTITASCPMDLQYFPMRDLCHIEI-----SFGYTRDIRYKNNEGPN 206
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
224 AVQVAGLTLPOFILRDEKDLG---CCTKHYNTEGKFTICIEVKFHELEOMGYLLQNTIP 279
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
207 SVGSSEVSLDQF-----KVLGHORAMEISLTGNYNSRLACEIOFVSMGYLLQIYIP 261
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
280 SLIYLISVSWFIMMDAPARVAGITVTLMTTQSSGSRASLPKVSVAIDIMAVCLLE 339
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
262 SGLIYIISVSWFIMMDAPARVAGITVTLMTTLMSTNAALPKISYVSIDVYLGTCF 321
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
340 LLEFVAALLEVAAINPVSQKKEITRLRRRQRLQLEEDTIOESRPFYRGYGLGHCLQAR 399
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
322 FVMVFPASLLEVATVGYMAK-----RIQMKQRFPAVQKMAEKKMQI----- 363
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
400 DGGPMEGSGIYSP-----QPPAP 417
|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```



Db 322 FVWFTSLLEYATVGYMAK-----R1QMRKORFTAVQKMOIDGPPGSAEPIPPRSTL 375

QY 379 -----IIQESRF-----YFRGYGLCHCQARDGPGMEGSGIYSPQPA 416

Db 376 SRPPSRLESEYRFVYHDPKAKSKGTLTENTINGARGP-----GPAIPA 419

RESULT 14

US-08-072-064-1

Sequence 1, Application US/08072064

Patent No. 6008046

GENERAL INFORMATION:

APPLICANT: FRENCH-CONSTANT, RICHARD H.

APPLICANT: JACKSON, MEYER B.

TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: PETER G. CARROLL

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/072.064

FILING DATE: 19930602

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 770,881

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: CARROLL, PETER G.

REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: OPHD-00574

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/705-8410

TELEFAX: 415/397-8338

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 637 amino acids

TYPE: AMINO ACID

TOPOLOGY: unknown

MOLECULE TYPE: peptide

ORGANISM: Drosophila melanogaster

POSITION IN GENOME:

CHROMOSOME/SEGMENT: III; polytene subregion 66F

MAP POSITION: approximately map unit 26

US-08-072-064-1

Query Match 32.0%; Score 720; DB 3; Length 637;

Best Local Similarity 36.5%; Pred. No. 1.4e-67;

Matches 155; Conservative 86; Mismatches 144; Indels 40; Gaps 9;

QY 5 VPATLSFLMLMTLPQGYLLRVALAKEE-----VKSGTGSGOPMSPSPDLKMGRTS 56

Db 46 LPRT-PLLTW-----LAIMMALIAOETGKRIHTVGAATGGGSLDGVNLSAIDLSFSV 99

QY 57 GYDARIRPNKGPVNVNTCNIFNSFSSVTKTMDRVNVNLFROQWMDPRLSYREYPD-D 115

Db 100 SYDKRVARNPGPVEVGVTMYVLSISSVSEVLMDFLDFYFRQFWTDPRLAYRKRPGVE 159

QY 116 SLDDPSMLDSIMKPDLPFANEKCANHEVTTDKLIRFKNGNVLYSIRLTLLSCIMD 175

Db 160 TLVSSEFIKINIMVDPFFVNEKOSYPHIATTSNEFIRVHNSGSITRSIRLTITASCPMN 219

QY 176 LKNRPMIOICTMQLJESSLISPLSLSVGTYTMDLVFEMLEDAFANOVAGLTLPO 235

Db 220 LQYFPMRQOLCHIEE-----SEGYTMDIRYFMRDGLGSVMGSSEVELQ 265

QY 236 FILDENDDLGCTTHYNTGKTCLIEVKFHLERONGYLLIQWIFSLILVLSWSEFINM 295

Db 266 FVILGHNO-RATEINLTGNTSRLACELOFVRSMGYLLIQIYISGLIYISWSEFINR 324

QY 296 DAAPARVGLITVLTMTTOSGSRASLPKSYVKAIDIMMAVCLLPFALLEYAINE 355

Db 325 NATPARVALGTYVLTMTLMSSTNALPKISYKSIDVYIGTCFVAVFASLEYATVGY 384

QY 356 VSRQKKEFIRLRORRORLEEDIIQESRFYRGYGLCHC---LQARDGPGMEGSGIYSP 412

Db 385 MAX-----R1QMRKORFMAIQIAEQKQOLDGANOQANPNNAVGP-GGVGYPG 437

QY 413 QPPAP 417

Db 438 GPGCP 442

RESULT 15

US-08-072-064-4

Sequence 4, Application US/08072064

Patent No. 6008046

GENERAL INFORMATION:

APPLICANT: FRENCH-CONSTANT, RICHARD H.

APPLICANT: JACKSON, MEYER B.

TITLE OF INVENTION: DRUG AND PESTICIDE SCREENING

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: PETER G. CARROLL

STREET: 220 Montgomery Street, Suite 2200

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/072.064

FILING DATE: 19930602

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 770,881

FILING DATE: 04-OCT-1991

ATTORNEY/AGENT INFORMATION:

NAME: CARROLL, PETER G.

REGISTRATION NUMBER: 32,837

REFERENCE/DOCKET NUMBER: OPHD-00574

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/705-8410

TELEFAX: 415/397-8338

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 637 amino acids

TYPE: AMINO ACID

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-072-064-4

Query Match 32.0%; Score 720; DB 3; Length 637;

Best Local Similarity 36.5%; Pred. No. 1.4e-67;

Matches 155; Conservative 86; Mismatches 144; Indels 40; Gaps 9;

QY 5 VPATLSFLMLMTLPQGYLLRVALAKEE-----VKSGTGSGOPMSPSPDLKMGRTS 56

Db 46 LPRT-PLLTW-----LAIMMALIAOETGKRIHTVGAATGGGSLDGVNLSAIDLSFSV 99

QY 57 GYDARIRPNKGPVNVNTCNIFNSFSSVTKTMDRVNVNLFROQWMDPRLSYREYPD-D 115

Db 100 SYDKRVARNPGPVEVGVTMYVLSISSVSEVLMDFLDFYFRQFWTDPRLAYRKRPGVE 159





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:55 ; Search time 28.9262 Seconds  
(without alignments)  
1432.403 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252

Sequence: 1 MTLVPATLSFLMLTLPGQ.....PQPPAPLLRCEGTRKLYVD 431

Scoring table:

BLAST62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1687	74.9	337	2	A49970 glycine receptor a
2	1682.5	74.7	452	2	S12381 glycine receptor a
3	1677	74.5	452	2	S14816 glycine receptor a
4	1671	74.2	452	2	S18836 glycine receptor a
5	1656.5	73.6	451	2	B49970 glycine receptor a
6	1592	70.7	464	2	A23682 glycine receptor a
7	1587.5	70.5	451	2	C49970 glycine receptor a
8	1580.5	70.2	449	2	S12382 glycine receptor a
9	1568.5	69.6	457	2	S20662 glycine receptor a
10	871	38.7	496	2	S46459 glycine receptor b
11	863.5	38.3	496	2	JH0165 glycine receptor b
12	855.5	38.0	497	2	G02031 glycine receptor b
13	742	32.9	499	2	S17785 gamma-aminobutyric
14	726	32.2	488	2	JH0359 gamma-aminobutyric
15	724	32.1	533	2	S33744 cyclodiene insecti
16	717	31.8	496	2	JH0603 gamma-aminobutyric
17	716	31.8	496	2	JH0828 gamma-aminobutyric
18	712	31.6	474	2	S04465 gamma-aminobutyric
19	712	31.6	474	2	S53531 gamma-aminobutyric
20	712	31.6	512	2	I52656 gamma-aminobutyric
21	706	31.3	474	2	JH0829 gamma-aminobutyric
22	705.5	31.3	1106	2	T25065 hypothetical prote
23	700.5	31.1	478	2	T24805 hypothetical prote
24	697	31.0	474	2	S53530 gamma-aminobutyric
25	696	30.9	473	2	A55275 gamma-aminobutyric
26	695.5	30.9	474	2	B60039 gamma-aminobutyric
27	694.5	30.8	474	2	B27142 gamma-aminobutyric
28	692.5	30.8	474	2	A40336 gamma-aminobutyric
29	692.5	30.8	654	2	C87791 protein B0207.12 f

30	692	30.7	473	2	S53532 gamma-aminobutyric
31	690	30.6	473	2	S04466 gamma-aminobutyric
32	687.5	30.5	476	2	S11440 gamma-aminobutyric
33	687	30.5	437	2	S34469 gamma-aminobutyric
34	673	29.9	434	2	S50865 gamma-aminobutyric
35	670.5	29.8	461	2	S50864 gamma-aminobutyric
36	668	29.7	606	2	A41145 gamma-aminobutyric
37	664.5	29.5	449	2	A34625 gamma-aminobutyric
38	653.5	29.0	449	2	A36303 gamma-aminobutyric
39	653.5	29.0	495	2	T20754 hypothetical prote
40	648.5	28.8	449	2	I2630 GABA receptor del
41	646.5	28.7	465	2	A38079 gamma-aminobutyric
42	643.5	28.6	545	2	T27614 hypothetical prote
43	634.5	28.2	453	2	S11087 gamma-aminobutyric
44	634	28.2	474	2	S13086 gamma-aminobutyric
45	631.5	28.0	473	2	A38627 gamma-aminobutyric

#### ALIGNMENTS

RESULT 1  
A49970  
glycine receptor alpha-4 chain - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: A49970  
R:Matzenbach, B.; Maulet, Y.; Sefton, L.; Courtler, B.; Avner, P.; Guenet, J.L.; Betz  
U. Biol. Chem. 269, 2607-2612, 1994  
A:Title: Structural analysis of mouse glycine receptor alpha subunit genes. Identiflc  
A:Reference number: A49970; MUID:94132024; PMID:7507926  
A:Accession: A49970  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: DNA  
A:Residues: 1-337 <MAT>  
A:Cross-references: GB:X75850; NID:9435513; PIDN:CAA53468.1; PID:9817957  
A:Gene: Glad  
A:Superfamily: acetylcholine receptor  
C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match	74.9%	Score 1687	DB 2	Length 337
Best Local Similarity	92.3%	Pred. No. 3.7e-137		
Matches 324	Conservative 7	Mismatches 6	Indels 14	Gaps 1
QY	25	VALAKEEYKSGTSGQSPSPDLKMGRTSGYDARIRPNKGGPVAVTCIFINFS	SS 84	
DB	1	VALAKEEYKSGTSGQSPSPDLKMGRTSGYDARIRPNKGGPVAVTCIFINFS	SS 60	
QY	85	VTKTMDYRVNVEFLRQNNDDPLSYREYDDSLDLPMSLDSIMKPDLEFANEKANFHE	144	
DB	61	VTEITMDYRVNVEFLRQNNDDPLSYREYDDSLDLPMSLDSIMKPDLEFANEKANFHE	120	
QY	145	VTTDNKLRIRFKNGVLSIRLTILSCLMDLKNPMDIOTCTMLESSTICSPLS	204	
DB	121	VTTDNKLRIRFKNGVLSIRLTILSCLMDLKNPMDIOTCTMLESSTICSPLS	167	
QY	205	LSVGTMDYRVNVEFLRQNNDDPLSYREYDDSLDLPMSLDSIMKPDLEFANEKANFHE	264	
DB	168	SEGVTMDYRVNVEFLRQNNDDPLSYREYDDSLDLPMSLDSIMKPDLEFANEKANFHE	226	
QY	265	LEROMGYVLIQMYIFSLILVILSNVSPFNTMDAPARVGLITVLTMTTOSSGSRASLP	324	
DB	227	LEROMGYVLIQMYIFSLILVILSNVSPFNTMDAPARVGLITVLTMTTOSSGSRASLP	286	
QY	335	KVSYKKAIDIMAVCLLEFVFAALLEYAAINVEYSRHKETIRLRORRORL	375	
DB	287	KVSYKKAIDIMAVCLLEFVFAALLEYAAINVEYSRHKETIRLRORRORL	337	
RESULT 2				
S12381				
glycine receptor alpha-2 chain - human				

C:Species: Homosapiens (man)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
C:Accession: S12381  
R:Grenningloh, G.; Schmieden, V.; Schofield, P.R.; Seeburg, P.H.; Siddique, T.; Mohandas  
EMBO J. 9, 771-776, 1990  
A:title: Alpha subunit variants of the human glycine receptor: primary structures, function  
A:Reference number: S12381; MUID:90183975; PMID:2155780  
A:Accession: S12381  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-452 <GR>  
A:Cross-references: GB:X52008; NID:gj31848; PIDN:CAA36257.1; PID:gj31849  
A>Note: The sequence from Fig. 2 is inconsistent with that from Fig. 1 in lacking 16-Phe  
C:Superfamily: acetylcholine receptor  
C:Keywords: transmembrane protein

Query Match 74.7%; Score 1682.5; DB 2; Length 452;  
Best Local Similarity 79.0%; Pred. No. 1,3e-16;  
Matches 324; Conservative 29; Mismatches 36; Indels 21; Gaps 5;

D 24 RVALAKE-EVSGYTKGSQPMSPDFDLKMGRTSGYDARIRPNFKGPVNTCNIFNSF 82  
| | :  
23 RTACCKOHDSSNGQPSTLSPDFDLKMRTSGYDARIRPNFKGPVNTCNIFNSF 82  
| | :  
OY 83 SSVTKTMDRVNVFLRQNMNDPRLSYREYPDDSLDLPMSLDISIKRPDLFEANEKANF 142  
| | :  
Db 83 GSVEETTMDEVNFVNFELRQNMNDSRLAYSEYPDSDLDPMSLDISIKRPDLFEANEKANF 142  
| | :  
OY 143 HEVTTDNKLIFPFNGNVLYSIRLTLLISCMDLNKNPMDIQCTOMOLESSIICPLPS 202  
| | :  
Db 143 HDVTTDNKLILSKNGKVLYSIRLTLLTSCMDLNKNPMDIQCTOMOLE----- 191  
| | :  
OY 203 LLSLVGYTMMKDLVEWMLADAPAVOVBGLTLPEILRDEKDLCGCTHYNTGKFTCIENVK 262  
| | :  
Db 192 ---SFGYTMNDLIREFMISDGP-VQVAEGLTLPGCLKEEKELGCTKHYNTGKFTCIENVK 247  
| | :  
OY 263 FHLRQMKXYLIOWYIPSLIVLILSWVSFWINMDAAPARVGITTYLTMTTOSSGSRAS 322  
| | :  
Db 248 FHLEOKMYLIOWYIPSLIVLILSWVSFWINMDAAPARVALGITTYLTMTTOSSGSRAS 307  
| | :  
OY 323 LPKSVYKAIDIMNAVCLLFVFALLLEYAALIEFAALIEFVSROHKEPIRRRRQRORLEEDIOE 382  
| | :  
Db 308 LPKSVYKAIDIMNAVCLLFVFALLLEYAALIEFAALIEFVSROHKEPIRRRRQRORLEEDYTRE 367  
| | :  
OY 383 SRFFPRGIGLHCICQARDGGPMESGIS-FQPAAPLRBEETTRKLYVD 431  
| | :  
Db 368 SRFFPSGVGMGHCUVRKATPANPLRPPQ---KDDDAIKKKFVD 413  
| | :

RESULT 3  
S14816  
glycine receptor alpha-2 chain variant A precursor - rat  
N:Alternate names: neonatal glycine receptor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: S14816; S15776; JN0112  
R:Akagi, H.; Hirai, K.; Hisanuma, F.  
FEBS Lett. 281, 160-166, 1991  
A:title: Cloning of a glycine receptor subtype expressed in rat brain and spinal cord d  
A:Reference number: S14816; MUID:91200276; PMID:1707830  
A:Accession: S14816  
A:Molecule type: DNA  
A:Residues: 1-452 <AK>  
A:Cross-references: EMBL:X57281; NID:g56743; PIDN:CAA0549.1; PID:g56744  
R:Kunse, J.; Kurtyov, A.; Maulec, Y.; Matosio, M.L.; Schmieden, V.; Betz, H.  
FEBS Lett. 283, 73-77, 1991  
A:title: Alternative splicing generates two isoforms of the alpha-2 subunit of the inh  
A:Reference number: S15776; MUID:91243883; PMID:1645300  
A:Accession: S15776  
A>Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-452 <KU>  
A:Cross-references: GB:X61159; NID:g288344; PIDN:CAA43471.1; PID:g288345

R:Kuhse, J.; Schmieden, V.; Betz, H.  
Neuron 5, 867-873, 1990  
A>Title: A single amino acid exchange alters the pharmacology of neonatal rat glycine receptors  
A:Reference number: JN0112; MUID:91097798; PMID:2176511  
A:Accession: JN0112  
A:Molecule type: mRNA  
A:Residues: 1-193, 'E', 195-452 <KU2>  
C:Comment: Glycine reduces neuronal firing by activating this inhibitory glycine receptor  
C:Superfamily: acetylcholine receptor  
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; transmembrane protein  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:128-452/Product: glycine receptor alpha-2 chain #status predicted <MAT>  
F:253-280/Domain: transmembrane #status predicted <TM2>  
F:287-304/Domain: transmembrane #status predicted <TM3>  
F:316-342/Domain: transmembrane #status predicted <TM4>  
F:424-441/Domain: transmembrane #status predicted <TM5>  
F:72,103/Binding site: carboxylate (asn) (covalent) #status predicted

Query Match: 74.5%; Score 1677; Db:2; Length 452;  
Best Local Similarity 75.5%; Pred. No. 3,96-116;  
Matches 327; Conservative 32; Mismatches 48; Indels 26; Gaps 6;

OY 1 MTLVPLATSLFLMLTLPQVLLKVALAKE-EVKSSTGRKSGQPMSPDPLDKMGRTSGYD 59  
Db 5 LVNLTLLFAEFL-----GTNHFREARCKDHSRSGKNHQSOTLSPSPDLKMGRTSGYD 59  
OY 60 ARIKPNFGPVPVNTCNFIINSFSSVTKTMDYRVANFRLROONDPRLSYREYDDSLD 119  
Db 60 ARIKPNFGPVPVNTCNFIINSFSGYVETTMIDYRVNFIQLQONHDSRLAISEYDDSLD 119  
OY 120 DPSMLDSIWKPDLEFPANKEGANFHEVTTDNKLLIFKNGVLYSIRLLTSLCLMDLKNF 179  
Db 120 DPSMLDSIWKPDLEFPANKEGANFHDVTTDNKLLISKNGVLYSIRLLTSLCPMDLKNF 179  
OY 180 PMDIOTGCMOLESSILICSLPLSLSLSGVYMKRLVPEMLDPAVAVAGTLTPOFILR 239  
Db 180 PMDIOTGCMOLE-----SEGYTMNDLIREMSDGP-VVVAEGLTLPQFILK 224  
OY 240 DEKDLGCTKHYNTGKFTCLEVKFHLEROMQGYLLIOMYIPSLILVILISVSWFWINMDAP 299  
Db 225 EEKEELGCTKHYNTGKFTCLEVKFHLEROMQGYLLIOMYIPSLILVILISVSWFWINMDAP 284  
OY 300 ARVGLGTTVLTMTTQSSGSRASLPKYSYVAIDIMMAVCLLPFAALLLEAATNPFVSRQ 359  
Db 285 ARVAGLTITVLTMTTQSSGSRASLPKYSYVAIDIMMAVCLLPFAALLLEAANPFVSRQ 344  
OY 360 HKEEIRLRRORRORLEEDITQESRPFYRGYGLCHCLQARDGPGMEGSGIYS-POPAPL 418  
Db 345 HKEEIRLRRORRORLEEDITQESRPFYRGYGLCHCLQARDGPGMEGSGIYS-POPAPL 401  
OY 419 LREGETTRKLYVD 431  
Db 402 KDADAIKKRFVD 413

RESULT 4  
S18836  
glycine receptor alpha-2 chain precursor variant B - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 21-Aug-1998  
C:Accession: S18836  
R:Kuhse, J.; Kuraytov, A.; Maulet, Y.; Malosio, M.L.; Schmieden, V.; Betz, H.  
FEBS Lett. 283, 73-77, 1991  
A>Title: Alternative splicing generates two isoforms of the alpha-2 subunit of the neuronal glycine receptor  
A:Reference number: S15776; MUID:91243883; PMID:1645300  
A:Accession: S18836  
A:Molecule type: mRNA  
A:Status: preliminary  
A:Residues: 1-452 <KUH>  
C:Superfamily: acetylcholine receptor  
C:Keywords: glycoprotein; membrane protein  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:128-452/Product: glycine receptor alpha-2 chain variant B #status predicted <MAT>

F:103/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.2%; Score 1671; DB 2; Length 452;

Best Local Similarity 75.1%; Pred. No. 1.3e-135; Mismatches 325; Conservative 33; Mismatches 49; Indels 26; Gaps 6;

```
QY 1 MTLVPAATSLFLMLTLPQVLLRVALAKE-EVKSCTKGSQPMSPDFLDKMGRTSGYD 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 LVNLTALFAEFL-----GTNHFREAFCKDHDSSGKHPSQTLSPDFLDKMGRTSGYD 59

QY 60 ARIRPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVLFRQNMNDPRLSYREYDDSDL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 ARIRPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVLFRQNMNDPRLSYREYDDSDL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 DPSMLDSIMKPDLEFFANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNF 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 DPSMLDSIMKPDLEFFANEKANFHEVTTDNKLRISKNKGLYSIRLTLSCLMDLKNF 179

QY 180 PMDIQTCTMOLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOFLR 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 PMDIQTCTMOLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOFLR 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 DEKDLGCTKHYNKGFCTIEVKFHLEROMGYLIOMYIPSLILVILSWSFWMMDAAP 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 225 EEKELGCTKHYNKGFCTIEVKFHLEROMGYLIOMYIPSLILVILSWSFWMMDAAP 284

QY 300 ARVGLGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFVFAALLEYAAINFSRQ 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 285 ARVALGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFVFAALLEYAAINFSRQ 344

QY 360 HKEFIRLRROKROKLEEDIIQESRFYFRGYGLCHCLOARDGPMESGIYS-PQPAFL 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 HKEFIRLRROKROKLEEDYRESRFNFSGYMGHCLOKDVATVAKTAPANPLPOP- 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 419 LREGETTRKLYVD 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 -KDADAIRKKFVD 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 5

B49970

glycine receptor alpha-2 chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-1999

C:Accession: B49970; S15802

R:Matzenbach, B.; Maulet, Y.; Sefton, L.; Courtier, B.; Avner, P.; Guenet, J.L.; Betz, H.

J. Biol. Chem. 269, 2607-2612, 1994

A:Title: Structural analysis of mouse glycine receptor alpha subunit genes. Identification

A:Reference number: A49970; MUID:94132024; PMID:7507926

A:Accession: B49970

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-451 <MAP>

A:Cross-references: GB:X75841

R:Kunze, J.; Kurtylow, A.; Maulet, Y.; Malosio, M.L.; Schmieden, V.; Betz, H.

FEBS Lett. 283, 73-77, 1991

A:Title: Alternative splicing generates two isoforms of the alpha-2 subunit of the inhibi

A:Reference number: S15776; MUID:91243883; PMID:1645300

A:Accession: S15802

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 67-89 <KUH>

C:Genetics:

A:Gene: Glra2

C:Superfamily: acetylcholine receptor

C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match 73.6%; Score 1656.5; DB 2; Length 451;

Best Local Similarity 75.1%; Pred. No. 2.2e-134;

Matches 325; Conservative 33; Mismatches 48; Indels 27; Gaps 7;

```
QY 1 MTLVPAATSLFLMLTLPQVLLRVALAKE-EVKSCTKGSQPMSPDFLDKMGRTSGYD 59
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 LVNLTALFAEFL-----GTNHFREAFCKDHDSSGKHPSQTLSPDFLDKMGRTSGYD 59
```

```
Db 5 LVNLTALFAEFL-----GTNHFREAFCKDHDSSGKHPSQTLSPDFLDKMGRTSGY - 58
QY 60 ARIRPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVLFRQNMNDPRLSYREYDDSDL 119
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 59 ARIRPNKGPVAVNTCNIFINSFSSVTKTMDYRVNVLFRQNMNDPRLSYREYDDSDL 118

QY 120 DPSMLDSIMKPDLEFFANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNF 179
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 DPSMLDSIMKPDLEFFANEKANFHEVTTDNKLRISKNKGLYSIRLTLSCLMDLKNF 178

QY 180 PMDIQTCTMOLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOFLR 239
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 PMDIQTCTMOLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOFLR 223
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 240 DEKDLGCTKHYNKGFCTIEVKFHLEROMGYLIOMYIPSLILVILSWSFWMMDAAP 299
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 224 EEKELGCTKHYNKGFCTIEVKFHLEROMGYLIOMYIPSLILVILSWSFWMMDAAP 283

QY 300 ARVGLGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFVFAALLEYAAINFSRQ 359
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 284 ARVALGITVLTMTTQSSGRASLPKVSYYKAIDIMNAVCLLFVFAALLEYAAINFSRQ 343

QY 360 HKEFIRLRROKROKLEEDIIQESRFYFRGYGLCHCLOARDGPMESGIYS-PQPAFL 418
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 344 HKEFIRLRROKROKLEEDYRESRFNFSGYMGHCLOKDVATVAKTAPANPLPOP- 400
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 419 LREGETTRKLYVD 431
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 401 -KDADAIRKKFVD 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 6

A23682

glycine receptor alpha-3 chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 20-Aug-1999

C:Accession: A23682

R:Kunze, J.; Schmieden, V.; Betz, H.

J. Biol. Chem. 265, 22317-22320, 1990

A:Title: Identification and functional expression of a novel ligand binding subunit o

A:Reference number: A23682; MUID:91093073; PMID:2176214

A:Accession: A23682

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-464 <KUH>

A:Cross-references: GB:M5250; GB:M38385; NID:q204882; PIDN:AAA63492.1; PID:q204883

C:Superfamily: acetylcholine receptor

C:Keywords: transmembrane protein

Query Match 70.7%; Score 1592; DB 2; Length 464;

Best Local Similarity 71.3%; Pred. No. 8e-129;

Matches 313; Conservative 34; Mismatches 50; Indels 42; Gaps 7;

```
QY 9 LSTLLMTLPQVLLRVALAKEEVKSGTKSQPMSPDFLDKMGRTSGYARIRPNFKG 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 LSGYFWE-----AALLSLVATKETSARSASPMSPDFLDKMGRTSGYARIRPNFKG 67

QY 69 PVAVNTCNIFINSFSSVTKTMDYRVNVLFRQNMNDPRLSYREYDDSDLDPMSLSIW 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 PVAVNTCNIFINSFSSVTKTMDYRVNVLFRQNMNDPRLSYREYDDSDLDPMSLSIW 127

QY 129 KPDLFFANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNFPMDIQTCTM 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 KPDLFFANEKANFHEVTTDNKLRIFKNGVLSIRLTLSCLMDLKNFPMDIQTCTM 187

QY 189 QLESSILCSPLPSLSVGYTKMDLVFEWLEDAVAOVAEGTLPOFLRDEKDLGCT 248
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 QLE-----SFGYTMNDLIFEWMDGP-VQVAGTLPLPOFLRDEKDLRQCT 232

QY 249 KHYNKGKTCIEVKFHLEROMGYLIOMYIPSLILVILSWSFWMMDAAPARVGLGTT 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 233 KHYNKGKTCIEVKFHLEROMGYLIOMYIPSLILVILSWSFWMMDAAPARVGLGTT 292
```

QY 309 VLTMTTSSGSRASLPKVSYSKAIIDIMAVCLLFVFAALLFVFAALINFEVSRQKEFIIRLR 368  
 DB 293 VLTMTTSSGSRASLPKVSYSKAIIDIMAVCLLFVFAALLFVFAALINFEVSRQKEFIIRLR 352  
 QY 369 RQRQ-----RLE-----EDIIQESRFYFGYGLGICLQARDG-----GPMESGSIYSP 412  
 DB 353 KKKNTTEAFALKEFYRSDTDEVRSHSLFTAYGKPCLOAKDGVKPKHNAVQV--- 409  
 QY 413 QPPAPLREGGTRKLYVD 431  
 DB 410 -----MPKSADMKRVFID 423

## RESULT 7

C49970  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Nov-1995 #sequence\_revision 17-Nov-1995 #text\_change 17-Mar-1999  
 C:Accession: C49970; S48662  
 R:Matzenbach, B.; Maulet, Y.; Sefton, L.; Courtier, B.; Avner, P.; Guenet, J.L.; Betz, H.  
 J. Biol. Chem. 269, 2607-2612, 1994  
 A:Title: Structural analysis of mouse glycine receptor alpha subunit genes. Identifica-  
 A:Reference number: A49970; MUID:94132024; PMID:7507926  
 A:Accession: C49970  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-451 <MAY>  
 A:Cross-references: GB:X75832  
 R:Saul, B.; Schmieden, V.; Kling, C.; Muelhardt, C.; Gass, P.; Kuhse, J.; Becker, C.M.  
 FEBS Lett. 350, 71-76, 1994  
 A:Title: Point mutation of glycine receptor alpha-1 subunit in the spasmodic mouse affect  
 A:Reference number: S48662; MUID:94341377; PMID:8062927  
 A:Accession: S48662  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 62-84 <SAU>  
 C:Genetics:  
 A:Gene: Glal  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: neurotransmitter receptor; transmembrane protein

Query Match 70.5%; Score 1587.5; DB 2; Length 451;  
 Best Local Similarity 72.8%; Pred. No. 1.9e-128;  
 Matches 310; Conservative 40; Mismatches 53; Indels 23; Gaps 7;  
 QY 8 TLSFLTLTGPQVLRVALAKEEVKSGTSGSPSPDFLDKLMGRTSGYDARIRPNFK 67  
 DB 6 TLRFLWETI--VFSLA-ASKEAARSAKPKSPDFLDKLMGRTSGYDARIRPNFK 61  
 QY 68 GPPAVATGNIPIINSSTYTKTMDYRVNVEFLRQONDRSLSTREYPPDSLDLPSMLDSI 127  
 DB 62 GPPAVATGNIPIINSSTYTKTMDYRVNVEFLRQONDRSLSTREYPPDSLDLPSMLDSI 121  
 QY 128 WKPDLEFANEGANHEVTTDNKILRFKNGNVLISIRLTLSCLMDLKNFPMIDICT 187  
 DB 122 WKPDLEFANEGANHEVTTDNKILRFKNGNVLISIRLTLSCLMDLKNFPMIDICT 181  
 QY 188 MOLESSITLSPSLSVGYTKMDLVEFWLEDAPAVQVAGTLTLPQITLDERDLOC 247  
 DB 182 MOLE-----SFQYTMNDLIFEWQEG-AYOVADGTLTLPQITLDERDLOC 226  
 QY 248 TKHYMTGKFTCLEVFNHERQMGYLLIOMYIPSLILVLSWSEFIMNDAPARVGLGIT 307  
 DB 227 TKHYMTGKFTCLEVFNHERQMGYLLIOMYIPSLILVLSWSEFIMNDAPARVGLGIT 286  
 QY 308 TLTMTTSSGSRASLPKVSYSKAIIDIMAVCLLFVFAALLFVFAALINFEVSRQKEFIIRLR 367  
 DB 287 TLTMTTSSGSRASLPKVSYSKAIIDIMAVCLLFVFAALLFVFAALINFEVSRQKEFIIRLR 346  
 QY 368 RQRQ-----RLE-----EDIIQESRFYFGYGLGICLQARDG-----GPMESGSIYSP--QPPAPLREGGTR 424  
 DB 347 RQRQ-----RLE-----EDIIQESRFYFGYGLGICLQARDG-----GPMESGSIYSP--QPPAPLREGGTR 405

QY 425 TRKLYV 430  
 DB 406 MKRLFI 411

## RESULT 8

S12382  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
 C:Accession: S12382  
 R:Grenningloh, G.; Schmieden, V.; Schofield, P.R.; Seeburg, P.H.; Stoddique, T.; Mohan  
 EMBO J. 9, 771-776, 1990  
 A:Title: Alpha subunit variants of the human glycine receptor: primary structures, fu  
 A:Reference number: S12381; MUID:90183975; PMID:2155780  
 A:Accession: S12382  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-449 <GR>  
 A:Cross-references: GB:X52009; NID:931850; PIDN:CAA36258.1; PID:931851  
 C:Superfamily: acetylcholine receptor  
 C:Keywords: transmembrane protein

Query Match 70.2%; Score 1580.5; DB 2; Length 449;  
 Best Local Similarity 73.4%; Pred. No. 7.5e-128;  
 Matches 307; Conservative 38; Mismatches 50; Indels 23; Gaps 7;

QY 17 LPGVLLRVALAKEEVKSGTSGSPSPDFLDKLMGRTSGYDARIRPNFKGPPVNTCN 76  
 DB 11 LSGAIVFSLASKEAEARSAKPKSPDFLDKLMGRTSGYDARIRPNFKGPPVNTCN 70  
 QY 77 IFINSFSSVTKTMDYRVNVEFLRQONDRSLSTREYPPDSLDLPSMLDSIMKPDLEFAN 136  
 DB 71 IFINSFSGIAETMDYRVNVEFLRQONDRSLSTREYPPDSLDLPSMLDSIMKPDLEFAN 130  
 QY 137 EKGANFHEVTTDNKILRFKNGNVLISIRLTLSCLMDLKNFPMIDICTMOLESSITL 196  
 DB 131 EKGANFHEVTTDNKILRFKNGNVLISIRLTLSCLMDLKNFPMIDICTMOLESSITL 185  
 QY 197 CSPLPSLSLVGYTKMDLVEFWLEDAPAVQVAGTLTLPQITLDERDLOCCTKHYNKGF 256  
 DB 186 -----SFQYTMNDLIFEWQEG-AYOVADGTLTLPQITLDERDLOCCTKHYNKGF 235  
 QY 257 TCIEVKFHLEROMGYLLIOMYIPSLILVLSWSEFIMNDAPARVGLGITVLTMTQS 316  
 DB 226 TCIEVKFHLEROMGYLLIOMYIPSLILVLSWSEFIMNDAPARVGLGITVLTMTQS 295  
 QY 317 SGRASLPLKVSYSKAIIDIMAVCLLFVFAALLFVFAALINFEVSRQKEFIIRLRQRORLE 376  
 DB 296 SGRASLPLKVSYSKAIIDIMAVCLLFVFAALLFVFAALINFEVSRQKEFIIRLRQRORLE 353  
 QY 377 EDIIQESRFYFGYGLG-HCLQARDGPMESGSIYSP--QPPAPLREGGTRKLYV 430  
 DB 354 EDEAGEGFNFSAVGMGPACLOAKDGVKPKHNAVQV---MPKSADMKRVFID 409

## RESULT 9

S20662  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 20-Aug-1999  
 C:Accession: S20662; A38597; A27141  
 R:Malosio, M.; Kuhse, J.; Betz, H.  
 submitted to the EMBL Data Library, December 1990  
 A:Description: Complete coding sequence of the an splicing variant of the rat glycine  
 A:Reference number: S20662  
 A:Accession: S20662  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-457 <MAL>  
 A:Cross-references: EMBL:X55246; NID:956468; PIDN:CAA3987.1; PID:956469  
 R:Malosio, M.L.; Grenningloh, G.; Kuhse, J.; Schmieden, V.; Stoddique, T.; Prior, P.; B  
 J. Biol. Chem. 266, 2048-2053, 1991









Tue Jul 1 10:05:51 2003

us-10-075-846-4.rpr

Page 8

```

Db      324 SIDVYLGICFVWVFVFSLEIYATVGMAKNIQIGKQRFMAIQKTAEOKKOQADANHPPP 363
Qy      378 -DIIQESRFYERGYELGH 394
          :      :      :      :
Db      384 PPVSDSHSHGHGSHGH 401

```

Search completed: June 25, 2003, 17:17:34  
Job time : 29.9262 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 25, 2003, 17:12:59 ; Search time 17.3557 Seconds

(without alignments)  
1029.995 Million cell updates/sec

Title: US-10-075-846-4

Perfect score: 2252  
Sequence: 1 MTTLVPAFLSFLMLTLPQ.....PPAPAPLREGETTRKLYVD 431

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1682.5	74.7	452	1	GRA2_HUMAN
2	1678	74.5	337	1	GRA4_MOUSE
3	1670	74.2	452	1	GRA2_RAT
4	1602.5	71.2	465	1	GRA3_HUMAN
5	1592	70.7	464	1	GRA3_RAT
6	1580.5	70.2	449	1	GRA1_HUMAN
7	1580	69.8	457	1	GRA1_BOVIN
8	1573	69.8	444	1	GRA1_BRARE
9	1569.5	69.7	457	1	GRA1_MOUSE
10	1568.5	69.6	457	1	GRA1_RAT
11	871	38.7	496	1	GRB_MOUSE
12	863.5	38.3	496	1	GRB_RAT
13	855.5	38.0	497	1	GRB_HUMAN
14	742	32.9	499	1	GAB_LYMT
15	740	32.2	488	1	GAB4_CHICK
16	720	32.0	496	1	GAB4_DROME
17	715	31.7	496	1	GAB3_DROME
18	712	31.6	474	1	GAB2_HUMAN
19	712	31.6	474	1	GAB2_MOUSE
20	697	31.0	474	1	GAB1_MOUSE
21	696	31.0	473	1	GAB1_HUMAN
22	695.5	30.9	474	1	GAB1_BOVIN
23	694.5	30.8	474	1	GAB1_MOUSE
24	692	30.7	473	1	GAB3_HUMAN
25	691.5	30.7	474	1	GAB3_MOUSE
26	687.5	30.5	476	1	GAB3_CHICK
27	676.5	29.7	452	1	GAB_HUMAN
28	669	29.0	638	1	GAAT_MOUSE
29	664.5	29.5	449	1	GAAT_RAT
30	663	29.4	440	1	GAAT_MOUSE
31	653.5	29.0	449	1	GAD_MOUSE
32	652.5	29.0	440	1	GAD_HUMAN
33	644.5	28.6	465	1	GAR2_HUMAN

34	640.5	28.4	453	1	GA6_HUMAN	Q16445	homo sapien
35	640	28.4	474	1	GAR1_MOUSE	P56475	mus musculus
36	638.5	28.4	453	1	GA6_MOUSE	P16305	mus musculus
37	637.5	28.3	632	1	GAAT_HUMAN	Q9un88	homo sapien
38	636	28.2	474	1	GAR1_RAT	P50572	rattus norv
39	634.5	28.2	453	1	GA6_RAT	P30191	rattus norv
40	634	28.2	474	1	GAC2_CHICK	P21548	gallus gall
41	633.5	28.1	465	1	GA6_CHICK	Q90845	gallus gall
42	631.5	28.0	473	1	GAR1_HUMAN	P24046	homo sapien
43	631.5	28.0	552	1	GA4_MOUSE	Q9d6f4	mus musculus
44	631	28.0	552	1	GA4_RAT	P28471	rattus norv
45	629	27.9	467	1	GAC3_MOUSE	P27681	mus musculus

## ALIGNMENTS

RESULT 1  
GRA2\_HUMAN STANDARD; PRT; 452 AA.  
ID P23416; 099862;  
AC 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Glycine receptor alpha-2 chain precursor.  
GN GLRA2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90183975; PubMed=2155780;  
RA Greengrath G., Schmieden V., Schofield P.R., Seebury P.H.,  
RT Stiddique T., Mohandas T.K., Becker C.M., Betz H.;  
RT Alpha subunit variants of the human glycine receptor: primary  
RT structures, functional expression and chromosomal localization of the  
RT corresponding genes.";  
RT EMBO J. 9:771-776(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98337669; PubMed=9674912;  
RA Cummings C.J., Dahle E.J.R., Zoghbi H.Y.;  
RT "Analysis of the genomic structure of the human glycine receptor  
RT alpha-2 subunit gene and exclusion of this gene as a candidate for  
RT Rett syndrome.";  
RL Am. J. Med. Genet. 78:176-178(1998).  
RN [3]  
RP SEQUENCE OF 1-360 FROM N.A.  
RX MEDLINE=97129407; PubMed=8973915;  
RA Monani U.R., Burghes A.H.M.;  
RT "Structure of the human alpha 2 subunit gene of the glycine receptor:  
RT use of vectorette and Alu-exon PCR.";  
RL Genome Res. 6:1200-1206(1996).  
CC - FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
CC NEURONAL FIRING)  
CC - SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)  
CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL  
CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE  
CC RECEPTOR CORE.  
CC - SUBCELLULAR LOCATION: Integral membrane protein.  
CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2\* (SHOWN HERE) AND ALPHA-  
CC 2B; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC - MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
CC - SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL:	X52008;	CAC36297.1; -.
DR	EMBL:	AF053495;	AAC35290.1; JOINED.
DR	EMBL:	AF053487;	AAC35290.1; JOINED.
DR	EMBL:	AF053488;	AAC35290.1; JOINED.
DR	EMBL:	AF053490;	AAC35290.1; JOINED.
DR	EMBL:	AF053491;	AAC35290.1; JOINED.
DR	EMBL:	AF053492;	AAC35290.1; JOINED.
DR	EMBL:	AF053493;	AAC35290.1; JOINED.
DR	EMBL:	AF053494;	AAC35290.1; JOINED.
DR	EMBL:	U77731;	AAB38272.1; -.
DR	EMBL:	U77724;	AAB38272.1; JOINED.
DR	EMBL:	U77725;	AAB38272.1; JOINED.
DR	EMBL:	U77726;	AAB38272.1; JOINED.
DR	EMBL:	U77727;	AAB38272.1; JOINED.
DR	EMBL:	U77728;	AAB38272.1; JOINED.
DR	EMBL:	U77729;	AAB38272.1; JOINED.
DR	EMBL:	U77730;	AAB38273.1; JOINED.
DR	EMBL:	U77731;	AAB38273.1; JOINED.
DR	EMBL:	U77732;	AAB38273.1; JOINED.
DR	EMBL:	U77725;	AAB38273.1; JOINED.
DR	EMBL:	U77726;	AAB38273.1; JOINED.
DR	EMBL:	U77727;	AAB38273.1; JOINED.
DR	EMBL:	U77728;	AAB38273.1; JOINED.
DR	EMBL:	U77729;	AAB38273.1; JOINED.
DR	EMBL:	U77730;	AAB38273.1; JOINED.
DR	Pfam:	S12381;	-.
DR	GeneW:	HGNC:4327;	GLRA2.
DR	MIM:	305990;	-.
DR	InterPro:	IPR000188;	GABA_receptor.
DR	InterPro:	IPR001175;	Neur_channel.
DR	Pfam:	PF02931;	Neur_chan_LBD.1.
DR	Pfam:	PF02932;	Neur_chan_memb.1.
DR	PRINTS:	PR00252;	NRIIONCHANNEL.
DR	TIGRFAMS:	TIGR00860;	LIC.1.
DR	PROSITE:	PS00236;	NEUROTR_ION_CHANNEL.1.
KW	Receptor,	Postsynaptic membrane;	Ionic channel; Glycoprotein; Signal transmembrane; Multigene family; Alternative splicing.
KM	POTENTIAL.		
FT	CHAIN	1	27
FT	DOMAIN	28	452
FT	TRANSMEM	28	253
FT	TRANSMEM	254	280
FT	TRANSMEM	287	304
FT	TRANSMEM	319	304
FT	DOMAIN	343	423
FT	TRANSMEM	424	441
FT	DISULFID	172	186
FT	DISULFID	232	243
FT	CARBOHYD	72	72
FT	CARBOHYD	103	103
FT	VARSPLIC	85	86
SO	SEQUENCE	452 AA;	52002 MW;
			VT --> IA (IN ISOFORM ALPHA-2B).
			9EET4B3F49A555EA CRC64;

```

Db      192  ---SFGYIMNDLIEFMLDGP-VQVAEGILTPQLFKEKEKELGCTKTHYNGKFCIEVK 247
QY      263  FHLEFQMGYYLIQMTIPSLIYILSWSEFWIMDAAPRVGLGTTVLTMTTQSSGRAS 322
Db      248  FHLEFQMGYYLIQMTIPSLIYILSWSEFWIMDAAPRVAGITTVLTMTTQSSGRAS 307
QY      323  LPKSVYVAIDIMMVAIVCLLEFPAALEFAATINFEVSRQKEFRRLRRRQROKLEDDIOE 382
Db      308  LPKSVYVAIDIMMVAIVCLLEFPAALEFAAVNFVSRQKEFRRLRRRQROKKEEDVIRE 367
QY      363  SREYRGYGLGHCLQARDGGEWGSIGYS-DPPAPRLREGETTAKLYVD 431
Db      368  SRFNPSGYGMGHCLQKDGDTAVKATPAMPDPQP-----KDGDAIKKKFVD 413

RESULT 2
GR44_MOUSE
ID      GRA4_MOUSE      STANDARD:      PRT;      337 AA.
AC      Q61603;
DT      15-DEC-1998 (Rel. 37, Created)
DT      15-DEC-1998 (Rel. 37, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Glycine receptor alpha-4 chain precursor (Fragment).
GN      GLRA4.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxId=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/c;
RX      MEDLINE=94132024; PubMed=7507926;
RA      Matzenbach B., Maulet Y., Sefton L., Courtlier B., Avner P.,
RA      Guenet J.-L., Betz H.,
RT      "Structural analysis of mouse glycine receptor alpha subunit genes.
RT      Identification and chromosomal localization of a novel variant.";
RL      J. Biol. Chem. 269:2607-2612(1994).
CC      -I- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTANSMITTER-GATED ION
CC      CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE
CC      CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF
CC      NEURONAL FIRING).
CC      -I- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)
CC      SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL
CC      MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE
CC      RECEPTOR CORE.
CC      -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -I- SIMILILANEOUS: THE ALPHA SUBUNIT BINDS STYCHINE.
CC      -I- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as they are not used in any way
CC      that would conflict with this statement and is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch)
CC
DR      EMBL; X75850; CAAS3468.1; -
DR      EMBL; X73851; CAAS3468.1; JOINED.
DR      EMBL; X75852; CAAS3468.1; JOINED.
DR      EMBL; X75853; CAAS3468.1; JOINED.
DR      MGD; MGI:95750; Glra4.
DR      InterPro: IPR000188; GABAA_receptor.
DR      InterPro: IPR001175; Neur_Channel.
DR      Pfam; PF02931; Neur_chan_LBD; 1.
DR      Pfam; PF02932; Neur_chan_memb; 1.
DR      TIGRFAMs; TIGR00860; Lic; 1.
DR      PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW      Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW      Transmembrane; Multigene family.
FT      NON_TER 1
FT      SIGNAL <1 10 POTENTIAL.

```

```

FT CHAIN 11 >337 GLYCINE RECEPTOR ALPHA-4 CHAIN.
FT DOMAIN 11 234 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 235 255 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT DISULFID 148 162 BY SIMILARITY.
FT DISULFID 209 220 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 337 337
SQ SEQUENCE 337 AA; 38774 MW; 9A3EDD7230578582 CRC64;

Query Match 74.58; Score 1678; DB 1; Length 337;
Best Local Similarity 92.08; Pred. No. 1.3e-135;
Matches 323; Conservative 7; Mismatches 7; Indels 14; Gaps 1;

OY 25 VALAKEEYKSGTSGSPSPDFDLKMGRTSGYDARIRPFKGPVAVTCNIFINSFSS 84
DB 1 VALAKEEYKSGKLSQSPSPDFDLKMGRTSGYDARIRPFKGPVAVTCNITINSFSS 60
OY 85 VTKTMDYRVNVFLRQOMNDPRLSYREXPDDSLDLPSMDSIMKPDLEFANEKANFHE 144
DB 61 VTEETMDYRVNVFLRQOMNDPRLAYREXPDDSLDLPSMDSIMKPDLEFANEKANFHE 120
OY 145 VTTDNKILRIFKKNVYSTRILITLISCLMDLKNFPHDIOTCTQMOLESSILCSPLPSLS 204
DB 121 VTTDNKILRIFKKNVYSTRILITLISCLMDLKNFPHDIOTCTQMOLESSILCSPLPSLS 167
OY 205 LSVGYTKMDLFEWLEDAVAQVAEGTLTPOFILRDEKDLGCTCKHYNTGKFCIEYKFH 264
DB 168 -SEGYTMDLFEWLEDAVAQVAEGTLTPOFILRDEKDLGCTCKHYNTGKFCIEYKFH 226
OY 265 LEROMGYLLIOMYIPSLIIVLISVSWFWINMDAARVAGLITTVLTWTQSSGSRASLP 324
DB 227 LEROMGYLLIOMYIPSLIIVLISVSWFWINMDAARVAGLITTVLTWTQSSGSRASLP 286
OY 325 KVSIVKADIDIMAVCLLFFPALLEYAAINFSVROHKEFTILRRORRORL 375
DB 287 KVSIVKADIDIMAVCLLFFPALLEYAAINFSVROHKEFTILRRORRORL 337

RESULT 3
GRA2_RAT STANDARD: PRT; 452 AA.
ID GRA2_RAT STANDARD: PRT; 452 AA.
AC P22771; G91W28;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycine receptor alpha-2* chain precursor (Neonatal isoform)
DE (Glycine receptor strychnine binding subunit).
GN GLRA2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF GLU-194.
RC TISSUE=Brain;
RX MEDLINE=91097798; PubMed=2176511;
RA Kuhse J., Schmieden V., Betz H.;
RT "A single amino acid exchange alters the pharmacology of neonatal rat
RT glycine receptor subunit.";
RL Neuron 5:867-873(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Spinal cord;
RX MEDLINE=91200276; PubMed=1707830;
RA Akagi H., Hirai K., Hishinuma F.;
RT "Cloning of a glycine receptor subtype expressed in rat brain and
RT spinal cord during a specific period of neuronal development.";
RL FEBS Lett. 281:160-166(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91243883; PubMed=1645300;

```

```

RA Kuhse J., Kuratov A., Maulet Y., Malosio M.L., Schmieden V.,
RA Betz H.;
RT "Alternative splicing generates two isoforms of the alpha 2 subunit
RT of the inhibitory glycine receptor.";
RL FEBS Lett. 283:73-77(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Beato M., Groot-Kormelinck P.J., Colquhoun D., Sivillotti L.G.;
RT "Concentration dependence of single channel currents through rat
RT recombinant alpha 1 glycine receptors.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION
CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE
CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF
CC NEURONAL FIRING).
CC -SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)
CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL
CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE
CC RECEPTOR CORE.
CC -SUBCELLULAR LOCATION: Integral membrane protein.
CC -ALTERNATIVE PRODUCTS: 2 ISOFORMS: ALPHA-2* (SHOWN HERE) AND ALPHA-
CC 2B. ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -TISSUE SPECIFICITY: THE GLYCINE RECEPTOR IS ABUNDANT IN THE SPINAL
CC CORD AND BRAINSTEM OF VERTEBRATES.
CC -DEVELOPMENTAL STAGE: THE ALPHA-2* SUBUNIT ISOFORM IS PRESENT ONLY
CC IN NEONATAL RATS. ISOFORMS ALPHA-2A AND ALPHA-2B ARE PRESENT IN
CC THE PRE- AND NEONATAL BRAIN. AT LATER POSTNATAL STAGES, ALPHA-2A
CC LEVELS GREATLY DECREASE WHILE ALPHA-2B IS BARELY DETECTABLE.
CC -MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.
CC -5 SUBSTITUTIONS: IDENTICAL TO THE HUMAN ALPHA-2 SUBUNIT, EXCEPT FOR
CC AT POSITION 194 (G -> E) ACCOUNTS FOR THE LOWER STRYCHNINE
CC SENSITIVITY OBSERVED IN NEONATAL RATS.
CC -SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL: X57281; CAA40549.1; -
DR EMBL: X61159; CAA43471.1; -
DR EMBL: AJ310837; CAC35981.1; -
DR PIR: JN0112; JN0112.
DR PIR: S14816; S14816.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_membr; 1.
DR PRINTS: PRO0252; NRIONCHANNEL.
DR TIGRFAMs: TIGR00860; LIGR_1.
DR PROSITE: PS00236; NRURTR_ION_CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family; Alternative splicing.
FT SIGNAL 1 27
FT CHAIN 28 452
FT DOMAIN 28 253 GLYCINE RECEPTOR ALPHA-2* CHAIN.
FT TRANSMEM 254 280 EXTRACELLULAR (PROBABLE).
FT TRANSMEM 287 304 PROBABLE.
FT TRANSMEM 319 342 PROBABLE.
FT DOMAIN 343 423 CYTOPLASMIC (PROBABLE).
FT TRANSMEM 424 441 PROBABLE.
FT SITE 194 194 IMPORTANT FOR GLYCINE AND STRYCHNINE
FT BINDING.
FT DISULFID 172 186 BY SIMILARITY.
FT DISULFID 232 243 BY SIMILARITY.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 85 86 VT -> 1A (IN ISOFORM ALPHA-2B).

```

FT VARIANT 194 194 E -> G.  
FT MUTAGEN 194 194 E->G: INCREASED SENSITIVITY TO  
FT STRYCHNINE.  
SQ SEQUENCE 452 AA: 52053 MW: A9BEC6D1552A3A1C CRC64:  
Query Match 74.2%; Score 1670; DB 1; Length 452;  
Best Local Similarity 75.3%; Pred. No. 9,3e-135;  
Matches 326; Conservative 32; Mismatches 49; Indels 26; Gaps 6;  
OY 1 MTLVPAFLSLFLMTLPQGVLRVALAKE-EVKSQTGSGSPSPDLKLMGRTSGVD 59  
DB 5 LVNITLALFAFL-----GTNHFRAFCCKDHSRSGKHSQTLSPDFLDKMGRTSGVD 59  
OY 60 ARIKPNKGPVNVTCNIFINSFSSVYKTTMDYRVNVELRQOMDPRLSREYPPDSIDL 119  
DB 60 ARIKPNKGPVNVTCNIFINSFSSVYKTTMDYRVNVELRQOMDPRLSREYPPDSIDL 119  
OY 120 DPMMDLSTKPDLEFANEKGANFHEVTDNKLIRFKGNVLYSTRLLTILSCMDLKNF 179  
DB 120 DPMMDLSTKPDLEFANEKGANFHEVTDNKLIRFKGNVLYSTRLLTILSCMDLKNF 179  
OY 180 PMDIQTCTMQLSESSILCSPLSLSVGYTKMDLVFEMLEDPAPVQVAGELTLPQFILR 239  
DB 180 PMDIQTCTMQLSESSILCSPLSLSVGYTKMDLVFEMLEDPAPVQVAGELTLPQFILR 239  
OY 240 DEKDLGCTKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIYLVSWSFWINMDAP 299  
DB 240 DEKDLGCTKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIYLVSWSFWINMDAP 299  
OY 225 EEKEKGYGCKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIYLVSWSFWINMDAP 284  
DB 225 EEKEKGYGCKHYNTGKFTCIKVEKFLEROMGYLLIQMTIPSLIYLVSWSFWINMDAP 284  
OY 300 ARVVGITTVLMTQSSGRASLPKVSYSKVIDIMAVCLFVFRALLEVAANFVSQ 359  
DB 285 ARVVGITTVLMTQSSGRASLPKVSYSKVIDIMAVCLFVFRALLEVAANFVSQ 344  
OY 360 HKEFTLRRLRRQRLEEDIIQESRFYRGYGLGCHLQARDGPEGSGIYS-POPPAPL 418  
DB 345 HKEFTLRRLRRQRLEEDIIQESRFYRGYGLGCHLQARDGPEGSGIYS-POPPAPL 401  
OY 419 LREGGTRKLYVD 431  
DB 402 -KDADAIRKKFYD 413  
RESULT 4  
GRA3\_HUMAN  
ID GRA3\_HUMAN STANDARD: PRT: 465 AA.  
AC 075311; 075816;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Glycine receptor alpha-3 chain precursor.  
GN GLRA3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=98344067; PubMed=9677400;  
RA Nikolich Z., Laube B., Weber R.G., Lichter P., Kioschis P., Poustka A.,  
RA Muehlhardt C., Becker C.-M.;  
RT "The human glycine receptor subunit alpha3. GLRA3 gene structure,  
RT chromosomal localization, and functional characterization of  
RT alternative transcripts.";  
RL J. Biol. Chem. 273:19708-19714(1998).  
CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
CC NEURONAL FIRING).  
CC -1- SUBUNIT: PENTAMER COMPOSED OF LIGAND BINDING ALPHA (48 kDa) AND  
CC STRUCTURAL BETA (58 kDa) SUBUNITS. A THIRD TYPE OF SUBUNIT (93  
CC kDa) IS A PERIPHERAL MEMBRANE PROTEIN ASSOCIATED WITH THE  
CC CYTOPLASMIC DOMAINS OF THE RECEPTOR CORE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ALPHA3L (SHOWN HERE) AND  
CC ALPHA3K: ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED THROUGHOUT THE CENTRAL  
CC NERVOUS SYSTEM.  
CC -1- DOMAIN: THE N-TERMINAL DOMAIN CARRIES STRUCTURAL DETERMINANTS  
CC ESSENTIAL FOR AGONIST AND ANTAGONIST BINDING. TM2 IS THOUGHT TO  
CC FORM THE INNER WALL OF THE CHLORIDE CHANNEL. THE CYTOPLASMIC LOOP  
CC IS AN IMPORTANT DETERMINANT OF CHANNEL INACTIVATION KINETICS.  
CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: AF017724; AAC39919.1; JOINED.  
CC EMBL: AF017715; AAC39919.1; JOINED.  
CC EMBL: AF017716; AAC39919.1; JOINED.  
CC EMBL: AF017717; AAC39919.1; JOINED.  
CC EMBL: AF017718; AAC39919.1; JOINED.  
CC EMBL: AF017719; AAC39919.1; JOINED.  
CC EMBL: AF017720; AAC39919.1; JOINED.  
CC EMBL: AF017721; AAC39919.1; JOINED.  
CC EMBL: AF017722; AAC39919.1; JOINED.  
CC EMBL: AF017723; AAC39919.1; JOINED.  
CC EMBL: U93917; AAC39917.1; JOINED.  
CC GeneW: HGNC:4328; GLRA3.  
CC MIM: 600421;  
CC InterPro: IPR000188; GABA\_A\_receptor.  
CC InterPro: IPR001175; Neur\_chan.  
CC Pfam: PF02932; Neur\_chan\_LBD; 1.  
CC Pfam: PF02932; Neur\_chan\_LBD; 1.  
CC DR PRINTS: PRO0252; NRIONCHANNEL.  
CC DR TIGRFAMs: TIGR00860; LIC; 1.  
CC DR PROSITE: PS00236; NEURORTR\_ION\_CHANNEL; 1.  
CC KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
CC Transmembrane; Multigene family; Alternative splicing.  
CC FT SIGNAL 1 33  
FT CHAIN 1 465  
FT DOMAIN 34 465 GLYCINE RECEPTOR ALPHA-3 CHAIN.  
FT TRANSMEM 253 279 EXTRACELLULAR (PROBABLE).  
FT TRANSMEM 286 303 PROBABLE.  
FT TRANSMEM 315 341 PROBABLE.  
FT DOMAIN 342 433 CYTOPLASMIC (PROBABLE).  
FT TRANSMEM 434 451 PROBABLE.  
FT DISULFID 171 185 BY SIMILARITY.  
FT DISULFID 231 242 BY SIMILARITY.  
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPPLIC 358 372 MISSING (IN ISOFORM ALPHA3K).  
FT CONFLICT 461 461 MISSING (IN REF. 1; AAC39917).  
SQ SEQUENCE 465 AA: 53937 MW: B8A267A310C6E64 CRC64:  
Query Match 71.2%; Score 1602.5; DB 1; Length 465;  
Best Local Similarity 73.8%; Pred. No. 5,5e-129;  
Matches 312; Conservative 35; Mismatches 43; Indels 33; Gaps 6;  
OY 22 LIRVALAKEVKSQTKGSPSPDLKLMGRTSGYDARIRPNKGPVNVTCNIFINS 81  
DB 21 LLLSLVATKETDSARSAPSPDLKLMGRTSGYDARIRPNKGPVNVTCNIFINS 80  
OY 82 ESSVYKTTMDYRVNVELRQOMDPRLSREYPPDSIDLDPMSLDSIMKPDLEFANEKAN 141  
DB 81 FIEVYTDNKLIRFKGNVLYSTRLLTILSCMDLKNFPMQVOTCTMQLSESSILCSPLP 201  
OY 142 FIEVYTDNKLIRFKGNVLYSTRLLTILSCMDLKNFPMQVOTCTMQLSESSILCSPLP 201  
DB 141 FIEVYTDNKLIRFKGNVLYSTRLLTILSCMDLKNFPMQVOTCTMQLSESSILCSPLP 190







```
FT FT VARIANT 299 299 /FTID-VAR_000298.  
FT FT VARIANT 304 304 R -> Q (IN STEE); DECREASED POTENCY OF  
FT FT VARIANT 307 307 /FTID-VAR_000299. GLYCINE TO ACTIVATE THE CHANNEL.  
FT FT VARIANT 307 307 /FTID-VAR_000300. K -> E (IN STEE).  
FT FT VARIANT 420 420 Y -> C (IN STEE).  
FT FT VARIANT 420 420 /FTID-VAR_000301. R -> H (IN STEE).  
FT FT SEQUENCE 449 AA: 51495 MW: 26653087B40918B CMC64; /FTID-VAR_010114.  
SO  
Query Match 70.2%; Score 1580.5; DB 1; Length 449;  
Best Local Similarity 73.4%; Pred. No. 3.9e-127;  
Matches 307; Conservative 38; Mismatches 50; Indels 23; Gaps 7;  
QY 17 LPGOVLLRALAKEEVKSGTSGSPSPDFLDKLGKRTSGYDARIRPNKGPVNTCN 76  
DB 11 LSGAIVFFSLAASKEAEARSATPMSPSDFLDKLGKRTSGYDARIRPNKGPVNTCN 70  
QY 77 IFINSFSSVTKTMDYRVNVLROQNNPRLSYREYPPDSDLDPMSMDLIMKPDLPFRAN 136  
DB 71 IFINSFSGIAETTMIDYRVNVLROQNNPRLAYNEYPDSDLDPMSMDLIMKPDLPFRAN 130  
QY 137 EKGANFHEVTDNKLRIFKNGNVLSIRLTLLSLMDLKNFPMIDLOTCTMOLESSIL 196  
DB 131 EKGAFHEITTDNKLIRSRNGNVLSIRLTLLACPMDLKNFPMIDVOTCTMOLE---- 185  
QY 197 CSPLPSLSLVGYTMKDLVFEWLEDAVAVOYAEGLTLPOFLIRDEKDLGCTTKHNTGKF 256  
DB 186 -----SFGYTMNDLIFEMOEG-AVOYADGLTLPOFLIRDEKDLGCTTKHNTGKF 235  
QY 257 TCIEVKFHLEROMGYLLIOMYIPSLIYLISMSFWMINDAARAVGLITVLTMTQS 316  
DB 236 TCIEARFHLEROMGYLLIOMYIPSLIYLISMSFWMINDAARAVGLITVLTMTQS 295  
QY 317 SGRSRASLPKVSYKAIIDIMAVCLLFVPAALLEYAALNFVSRQHKFIRLRORRORLE 376  
DB 296 SGRSRASLPKVSYKAIIDIMAVCLLFVPAALLEYAALNFVSRQHKFIRLRORRORLE 353  
QY 377 EDITOESFRFGYGLG-HCTQARDGGMEG---SGIYSPPPAPRLREGTTRKLYV 430  
DB 354 EDEAGEGRFNFSAYGMGPACLOAKDGISVKGANNSNTNP-PPAP-SKSPPEMRKLF 409  
RESULT 7  
GRAL_BOVIN STANDARD: PRT: 457 AA.  
AC P57695;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Glycine receptor alpha-1 chain precursor (Glycine receptor 48 kDa subunit) (Strychnine binding subunit).  
GN GLRA1.  
OS Bos taurus (Bovine).  
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; CC Bovidae; Bovinae; Bos.  
CC NCBI_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21109390; PubMed=11178872;  
RA Pierce K.D., Handford C.A., Morris R., Vafa B., Dennis J.A.,  
RA Healy P.J., Schofield P.R.;  
RT "A nonsense mutation in the alpha subunit of the inhibitory glycine receptor associated with bovine myoclonus".  
RL Mol. Cell. Neurosci. 17:354-363(2001).  
CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTANSITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING) (BY SIMILARITY).  
CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)
```

```
CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL  
CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE  
CC RECEPTOR CORE (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
CC -----  
DR EMBL: AF268375; AAG14346.1; -  
DR InterPro: IPR000188; GABA_A_receptor.  
DR InterPro: IPR001175; Neur_channel.  
DR Pfam: PF02931; Neur_chan_LBD; 1.  
DR Pfam: PF02932; Neur_chan_memb; 1.  
DR PRINTS: PR00252; NRIONCHANNEL.  
DR TIGRFS: TIGR00860; LTC; 1.  
DR PROSITE: PS00236; NEURON_ION_CHANNEL; 1.  
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;  
KW Transmembrane; Multigene family.  
FT SIGNAL 1 28 BY SIMILARITY.  
FT CHAIN 29 457 GLYCINE RECEPTOR ALPHA-1 CHAIN.  
FT DOMAIN 29 247 EXTRACELLULAR (PROBABLE).  
FT TRANSMEM 248 274 PROBABLE.  
FT TRANSMEM 281 298 PROBABLE.  
FT TRANSMEM 313 336 PROBABLE.  
FT DOMAIN 337 420 CYTOPLASMIC (PROBABLE).  
FT TRANSMEM 429 446 PROBABLE.  
FT DISULFD 166 180 BY SIMILARITY.  
FT DISULFD 226 237 BY SIMILARITY.  
FT CARBOHD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT BINDING 225 225 STRYCHNINE (OR 230) (BY SIMILARITY).  
FT BINDING 230 230 STRYCHNINE (OR 225) (BY SIMILARITY).  
SQ SEQUENCE 457 AA: 52613 MW: 0C60A814C8ABDDDE CMC64;  
Query Match 70.2%; Score 1580; DB 1; Length 457;  
Best Local Similarity 72.0%; Pred. No. 4.4e-127;  
Matches 309; Conservative 38; Mismatches 52; Indels 30; Gaps 8;  
QY 12 LLLMTLPQGVLLRALAKEEVKSGTSGSPSPDFLDKLGKRTSGYDARIRPNKGPV 71  
DB 9 LYLME--TIVFFSLAASKEAEARSASPMSPSDFLDKLGKRTSGYDARIRPNKGPV 65  
QY 72 NVTCNIFINSFSSVYKTTMDYRVNVLROQNNPRLSYREYPPDSDLDPMSMDLIMKPD 131  
DB 66 NVSCNIFINSFSGIAETTMIDYRVNVLROQNNPRLAYNEYPDSDLDPMSMDLIMKPD 125  
QY 132 LFPANKECANHEVTDNKLRIFKNGNVLSIRLTLLSLMDLKNFPMIDLOTCTMOLE 191  
DB 126 LFPANKECANHEVTDNKLRISNGNVLSIRLTLLACPMDLKNFPMIDVOTCTMOLE 185  
QY 192 SSSILCSPLPSLSLVGYTMKDLVFEWLEDAVAVOYAEGLTLPOFLIRDEKDLGCTTKH 251  
DB 186 -----SFGYTMNDLIFEMOEG-AVOYADGLTLPOFLIRDEKDLGCTTKH 230  
QY 252 NTGKFTCLEVKFHLEROMGYLLIOMYIPSLIYLISMSFWMINDAARAVGLITVLT 311  
DB 221 NTGKFTCLEARFHLEROMGYLLIOMYIPSLIYLISMSFWMINDAARAVGLITVLT 290  
QY 312 MTTQSSGSRASLPKVSYKAIIDIMAVCLLFVPAALLEYAALNFVSRQHKFIRLRORR 371  
DB 291 MTTQSSGSRASLPKVSYKAIIDIMAVCLLFVPAALLEYAALNFVSRQHKFIRLRORR 350  
QY 372 ROR-----LEEDITOESFRFGYGLG-HCTQARDGGMEG---SGIYSPPPAPRLLE 421  
DB 351 HHSKSMNLNLPDEDEAGEGRFNFSAYGMGPACLOAKDGISVKGANNSNTNP-PPAP-SKS 408  
QY 422 GETTRKLYV 430
```

DB 409 PEEMRLFT 417

RESULT 8

GRAL\_BRARE STANDARD: PRT: 444 AA.

ID GRAL\_BRARE 093430;

AC 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycine receptor, alpha1 chain precursor.

GN GLRA1.

OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

OC NCBI\_TaxId=7955;

OX NCBI\_TaxId=7955;

RN (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=99202839; PubMed=10189956;

RA David-Mattine B., Goblet C., de Saint Jan D., Fucile S., Devignot V., Bregestovski P., Korn H.;

RT Cloning, expression and electrophysiological characterization of glycine receptor alpha subunit from zebrafish.";

RL Neuroscience 90:303-317(1999).

CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.

CC -1- MISCELLANEOUS: THIS RECEPTOR IS ACTIVATED BY GLYCINE AND ANTAGONIZED BY STRYCHNINE. CAN ALSO BE ACTIVATED BY GABA AND INHIBITED BY BICUCULLINE.

CC -1- MISCELLANEOUS: HIGHLY SENSITIVE TO ACTIVATION BY TAURINE DESPITE THE PRESENCE OF A VALINE IN POSITION 135. IN MAMMALS VALINE AT THIS POSITION CAUSES A DRASTIC LOSS OF TAURINE EFFICACY.

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib.ch](mailto:license@sib.ch)).

CC -----

DR EMBL: AJ005812; CA06711.1; -

DR ZFIN: ZDB-GENE-991117-1; glra1.

DR InterPro: IPR000188; GABA\_A\_receptor.

DR InterPro: IPR001175; Neur\_channel.

DR Pfam: PF02931; Neur\_chan\_LBD; 1.

DR Pfam: PF02932; Neur\_chan\_memb; 1.

DR PRINTS: PR00252; NRIONCHANNEL.

DR TIGRPFAMs: TIGR00860; LIC; 1.

DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.

DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal; Transmembrane; Multigene family.

KW SIGNAL

FT 1 24

FT CHAIN 1

FT DOMAIN 25 444

FT TRANSMEM 248 268

FT TRANSMEM 277 294

FT TRANSMEM 311 331

FT DOMAIN 332 415

FT TRANSMEM 416 436

FT DISULFID 162 176

FT DISULFID 222 233

FT CARBOHYD 62 62

FT BINDING 221 221

FT BINDING 226 226

FT SIGNAL

FT BINDING 226 226

SQ SEQUENCE 444 AA: 50765 MW: 7766FC8C1A80581 CRC64;

Query Match 69.8%; Score 1573; DB 1: Length 444;

Best Local Similarity 71.9%; Pred. No. 1.7e-126;

Matches 302; Conservative 43; Mismatches 33; Indels 22; Gaps 5;

QY 12 LLLWTLPGQVLLRVALLAKEEVKSGTKGSGQSPSPDLKLMGRGTGYARIRPNFKGPPV 71

DB 6 IYLM-----ETIVFFSLAASQQAARKAASPMPSEFLDKLMKGSGYDARIRPNFKGPPV 61

QY 72 NVTGCFINSESSVYKTTMDYRVNVPFLRQOMDPRLSTREYVDDSLDPSMLDSIKMPD 131

DB 62 NVTGCFINSESSVYKTTMDYRVNVPFLRQOMDPRLSTREYVDDSLDPSMLDSIKMPD 121

QY 132 LFFANKEGNFHEVTTDNKLLRIFKNGVLYSIRLTLISCLMDKNFPMQDTQWOLE 191

DB 122 LFFANKEGNFHEVTTDNKLLRISKNGLVLSIRITVLACMDKNFPMQDTQWOLE 181

QY 192 SSSILCSPLPSLSLSGYTKMDLVEFMELEDAFAVAVAGTLTLPQFILRDEKDLGCTKHY 251

DB 182 -----SFGYTMNDLIEFW--DEKGAVOVAGSLTLPQFILRDEKDLGCTKHY 226

QY 252 NTGKFTCLEVKHLEBOKMGVYLQMYISLLIVILSWSEWIMDPAARVGLITVLT 311

DB 227 NTGKFTCLEARPHLEBOKMGVYLQMYISLLIVILSWSEWIMDPAARVGLITVLT 286

QY 312 MTTGSSGRASLPKYSYKATIDIMAVCLLPEFALLLEFAMINPFSRQHKERIRLROR 371

DB 287 MTTGSSGRASLPKYSYKATIDIMAVCLLPEFALLLEFAMINPFSRQHKERIRLROR --R 344

QY 372 RQLEEDIIQESRPFYFEGYGLG-HCLQARDGGPMGSGTSPQPPAPLREGETTRKLYV 430

DB 345 RRLKEDEAGDGRFSFAAVGMGPACIQAKDMAIKGNMNNAPTSTNPPEKTYEEMRKLEI 404

RESULT 9

GRAL\_MOUSE STANDARD: PRT: 457 AA.

ID GRAL\_MOUSE 064018; Q9R0Y6; Q9R0Y7;

AC 064018; Q64019; Q9R0Y6; Q9R0Y7;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycine receptor alpha-1 chain precursor (Glycine receptor 48 kDa subunit) (Strychnine binding subunit).

GN GLRA1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI\_TaxId=10090;

OX NCBI\_TaxId=10090;

RN (1)

RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND VARIANT SPD SER-80.

RX MEDLINE=95004575; PubMed=7920629;

RA Ryan S.G., Buckwalter M.S., Lynch J.W., Handford C.A., Segura L., Shing R., Wasmuth J.J., Camper S.A., Schofield P., O'Connell P.;

RT A missense mutation in the gene encoding the alpha 1 subunit of the inhibitory glycine receptor in the spasmodic mouse.";

RL Nat. Genet. 7:131-135(1994).

RP STRAIN=BALE/C.

RC MEDLINE=94132024; PubMed=7507926;

RX Matzenbach B., Maulet Y., Sefton L., Courtier B., Avner P., Guenet J.L., Betz H.;

RT "Structural analysis of mouse glycine receptor alpha subunit genes. Identification and chromosomal localization of a novel variant.";

RL J. Biol. Chem. 269:2607-2612(1994).

CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).

CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 KDA) AND BETA (58 KDA) SUBUNITS. A THIRD TYPE OF SUBUNIT (93 KDA) IS A PERIPHERAL MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE

CC RECEPTOR CORE.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A

CC SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- DISEASE: DEFECTS IN GLRA1 ARE A CAUSE OF SPASMODIC (SP), A MOUSE

CC MUTANT WHICH RESEMBLES TO THE HUMAN NEUROLOGICAL DISEASE,

CC HYPEREPILEPSIA (OR STARTLE DISEASE (STHE)).

CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.

CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: S73717; AAB32157.2; -

CC EMBL: S73718; AAB32158.2; -

CC EMBL: X75832; CAB52398.1; -

CC EMBL: X75833; CAB52398.1; JOINED.

CC EMBL: X75834; CAB52398.1; JOINED.

CC EMBL: X75835; CAB52398.1; JOINED.

CC EMBL: X75836; CAB52398.1; JOINED.

CC EMBL: X75837; CAB52398.1; JOINED.

CC EMBL: X75838; CAB52398.1; JOINED.

CC EMBL: X75839; CAB52398.1; JOINED.

CC EMBL: X75840; CAB52398.1; JOINED.

CC EMBL: X75832; CAB52399.1; -

CC EMBL: X75833; CAB52399.1; JOINED.

CC EMBL: X75834; CAB52399.1; JOINED.

CC EMBL: X75835; CAB52399.1; JOINED.

CC EMBL: X75836; CAB52399.1; JOINED.

CC EMBL: X75837; CAB52399.1; JOINED.

CC EMBL: X75838; CAB52399.1; JOINED.

CC EMBL: X75839; CAB52399.1; JOINED.

CC EMBL: X75840; CAB52399.1; JOINED.

CC MGD: MGI:95747; Glra1.

CC Interpro: IPR000188; GABA\_receptor.

CC Interpro: IPR001175; Neur\_channel.

CC Pfam: PF02931; Neur\_chan\_LBD.1.

CC Pfam: PF02932; Neur\_chan\_memb.1.

CC PRINTS: PR00252; NRIONCHANNEL.

CC TIGRFAMs: TIGR00860; LIC.1.

CC PROSITE: PS00236; NEUOTR\_ION\_CHANNEL.1.

CC Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;

CC Transmembrane; Multigene family; Alternative splicing;

CC Disease mutation.

CC SIGNAL 1 28

CC CHAIN 29 457 BY SIMILARITY.

CC DOMAIN 29 247 GLYCINE RECEPTOR ALPHA-1 CHAIN.

CC TRANSMEM 248 274 EXTRACELLULAR (PROBABLE).

CC TRANSMEM 281 298 PROBABLE.

CC TRANSMEM 313 336 PROBABLE.

CC DOMAIN 337 420 CYTOPLASMIC (PROBABLE).

CC TRANSMEM 429 446 PROBABLE.

CC DISULFID 166 180 BY SIMILARITY.

CC DISULFID 226 237 BY SIMILARITY.

CC CARBOHYD 66 66 N-LINKED (GLUCAC. . .) (PROBABLE).

CC BINDING 225 225 STRYCHNINE (OR 220) (BY SIMILARITY).

CC BINDING 230 230 STRYCHNINE (OR 225) (BY SIMILARITY).

CC VARSPLIC 354 361 MISSING (IN SHORT ISOFORM).

CC VARIANT 80 84 A -> S (IN SPD).

CC CONFLICT 84 84 M -> I (IN REF. 2).

CC CONFLICT 426 429 ISRI -> NISH (IN REF. 2).

CC SEQUENCE 457 AA: 52656 MW: 29268DC4991A6E20 CRC64;

Query Match 69.7%; Score 1569.5; DB 1; Length 457;

Best Local Similarity 71.5%; Pred. No. 3.5e-126;

Matches 309; Conservative 40; Mismatches 54; Indels 29; Gaps 8;

0y 8 TLSFLMLMTLPBGVLLNVALAKEEVKSGTGSGDPMSPSDFLDKLMGRTSGYDARIRPNFK 67

Db 6 TLFSEYLMETI---VFESLA-ASKEAEARSAAPKPMSPSDFLDKLMGRTSGYDARIRPNFK 61

Qy 68 GPPVNTVNTNFISFSFVTTTHDYRNVNVLROOANDPRLSTKEYEPDDSLDLPMSLDSI 127

Db 62 GPPVNTVNTNFISFSFVTTTHDYRNVNVLROOANDPRLSTKEYEPDDSLDLPMSLDSI 121

Qy 128 WKPDLPFANRKGANFHEVTVDNKLIRFRKNGVNLVSTRLLISCLMDLKNPMDIOTCT 187

Db 122 WKPDLPFANRKGAFHEITTDNKLIRSRGNNVLSTRITTLACPDMLKNPMDIOTCT 181

Qy 188 MOLESSIIICPLPSLSLVSQYTWKDLVFEMLEDPAPVQVAEGITLTPQFLRDEKLGCC 247

Db 182 MOLE-----SEGYTMNDLIFEMOEG-AYQVADGTLTPQFLRDEKLGCC 226

Qy 248 TKHYNTGKFTCIIEKFLEROMGYLLQWYIPSLIYIISWSPFWIMDAAPRVGIGT 307

Db 227 TKHYNTGKFTCIIEKFLEROMGYLLQWYIPSLIYIISWSPFWIMDAAPRVGIGT 286

Qy 308 TVLMTTQSSGRASLPKVSIVKAIDIMAVCLTFVFAALLEVAALNFVSRKHKEFTRLR 367

Db 287 TVLMTTQSSGRASLPKVSIVKAIDIMAVCLTFVFAALLEVAALNFVSRKHKEFTRLR 346

Qy 368 RQRROR-----LEEDIIDSRFERYGGLG-HCLQARDGGMESGIGTSP--QPPAPL 418

Db 347 RKRHHKSPMLNLFQDDEGEGERFNFSAYGMGPACLOAKDGI SVKGANNNNTTNPAP- 405

Qy 419 LREGETRKLIV 430

Db 406 SKSPEEMRKLFI 417

RESULT 10

GRAL\_RAT STANDARD; PRT; 457 AA.

ID GRAL\_RAT

AC P07727;

DT 01-APR-1988 (Rel. 07, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glycine receptor alpha-1 chain precursor (Glycine receptor 48 kDa subunit) (Strychnine binding subunit).

GN GLRA1 OR GLYR.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OC NCBI\_TaxID=10116;

OX [1]

RN SEQUENCE FROM N.A. (SHORT ISOFORM).

RX MEDLINE=91367372; PubMed=1716350;

RA Akagi H., Hirai K., Hishinuma F.;

RT "Functional properties of strychnine-sensitive glycine receptors expressed in *Xenopus* oocytes injected with a single mRNA.";

RL Neurosci. Res. 11:28-40(1991).

RN [2]

RP SEQUENCE OF 23-457 FROM N.A. (SHORT ISOFORM), AND SEQUENCE OF 29-49.

RC STRAIN=Wistar;

RX MEDLINE=87258250; PubMed=3037383;

RA Grenningloh G., Rindert A., Schmitt B., Methfessel C., Zensen M., Beyreuther K., Gundelfinger E.D., Betz H.;

RT "The strychnine-binding subunit of the glycine receptor shows homology with nicotinic acetylcholine receptors.";

RL Nature 328:215-220(1987).

RN [3]

RP SEQUENCE OF 1-30 FROM N.A.

RX MEDLINE=90180468; PubMed=2483325;

RA Sontheimer H., Becker C.M., Pritchett D.B., Schofield P.R., Greeningloh G., Kettemann H., Betz H., Seeburg P.H.;

RT "Functional chloride channels by mammalian cell expression of rat glycine receptor subunit.";

RL Neuron 2:1491-1497(1989).

RN [4]

RP STRYCHNINE BINDING SITE.

RX MEDLINE=91027731; PubMed=2171639;

RA Ruiz-Gomez A., Morato E., Garcia-Calvo M., Valdivieso F.,  
 RA Mayor F. Jr.;  
 RT "Localization of the strychnine binding site on the 48-kilodalton  
 RT subunit of the glycine receptor.";  
 RL Biochemistry 29:7033-7040(1990).  
 RN (5)  
 RP ALTERNATIVE SPLICING.  
 RC TISSUE-Spinal cord;  
 RX MEDLINE-9115810; PubMed-1703526;  
 RA Malosio M.L., Greeningloh G., Kuhse J., Schmieden V., Schmitt B.,  
 RA Prior P., Betz H.;  
 RT "Alternative splicing generates two variants of the alpha 1 subunit  
 RT of the inhibitory glycine receptor.";  
 RL J. Biol. Chem. 266:2048-2053(1991).  
 CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
 CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
 CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
 CC NEURONAL FIRING).  
 CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa)  
 CC SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL  
 CC MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE  
 CC RECEPTOR CORE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- MISCELLANEOUS: THE ALPHA SUBUNIT BINDS STRYCHNINE.  
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC  
 DR EMBL: D00833; BA00707.1; -  
 DR EMBL: Y00276; CAA68378.1; -  
 DR EMBL: M63915; AAA63490.1; -  
 DR EMBL: X55246; CAA38987.1; -  
 DR PIR: A27141; A27141.  
 DR PIR: A38597; A38597.  
 DR PIR: JN0014; JN0014.  
 DR InterPro: IPR000188; GABA\_A\_receptor.  
 DR InterPro: IPR001175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_memb; 1.  
 DR PRINTS: PR00257; NRIONCHANNEL.  
 DR TIGRFS: TIGR00860; LIC1.1.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 DR Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; signal;  
 KW Transmembrane; Multigene family; Alternative splicing.  
 FT SIGNAL 1 28  
 FT CHAIN 1 457  
 FT DOMAIN 29 247  
 FT TRANSMEM 248 274  
 FT TRANSMEM 281 298  
 FT TRANSMEM 313 336  
 FT DOMAIN 337 420  
 FT TRANSMEM 421 445  
 FT DISULFID 166 180  
 FT DISULFID 226 237  
 FT CARBOHYD 66 66  
 FT BINDING 225 225  
 FT BINDING 230 230  
 FT VARSPPLIC 354 361  
 FT CONFLICT 1 22  
 FT CONFLICT 26 26  
 FT CONFLICT A -> V (IN REF. 2).  
 SQ SEQUENCE 457 AA; 52616 MW; 0379C33A83C71FAA CRC64;

Matches 308; Conservative 39; Mismatches 57; Indels 27; Gaps 7;  
 QY 8 TLSELLMTLPGQVLLRYALAKEEYKSGTKSOPSPEDFLDKLMGRSGVDARIPRFK 67  
 DB 6 TLREYLMETI---VFSSLA-ASKEDAAKRSAPKPPSPDFDKLGRISGVDARIPRFK 61  
 QY 68 GPPVAVTCNIFINSFSSYTKTMDYRVNVFPLRQOQNDPRLSYREXPDDSLDPSMLDSI 127  
 DB 62 GPPVAVSCNIFINSFSGSLAETTMODRVNIFLRQOQNDPRLAYNRPDDSLDPSMLDSI 121  
 QY 128 WKPDLPFANEGANHEVTTDNKLRIRKNGNVLSIRLLTILSCMDLAKNFPMDIQCT 187  
 DB 122 WKPDLPFANEGANHEVTTDNKLRIRKNGNVLSIRLLTILACPMDLKNFPMDIQCTI 181  
 QY 168 MOLESSILCSPLSLSVGYTMDLVEFMLEDAQVAVQEGTLTPOFILRDKDGLCC 247  
 DB 162 MOLE-----SFGITMDLLEFMEQEG-AQVADGLTLPFILKEKDLRYC 226  
 QY 248 TKHYVTGKFTCLIEVNFHLEROMGYLLIOMYIPSLDIVILISVSEFMINNDAPARVGLGIT 307  
 DB 227 TKHYVTGKFTCLIEARFHLERQMGYLLIOMYIPSLDIVILISVSEFMINNDAPARVGLGIT 286  
 QY 308 TVLTMTTSSSGSRASLPRVSYVYKATDIMAVCLLFVFAALLEYAINFVSROHKEFTLR 367  
 DB 287 TVLTMTTSSSGSRASLPRVSYVYKATDIMAVCLLFVFAALLEYAVNFVSROHKEFTLR 346  
 QY 368 RRORQR-----LEEDTIOESRFYRGYGLG-HCLQARDGCPMEGSGISYPOPAFL-L 419  
 DB 347 RRRRHKSPMLNLPDDEGGEGRRFNFSAVGMCPACLOAKDGLSVGANNNTTNPAPAPS 406  
 QY 420 REGETTRKLYV 430  
 DB 407 KSPDEMRKLF 417  
 RESULT 11  
 GRB\_MOUSE  
 ID GRB\_MOUSE STANDARD: PRT: 496 AA.  
 AC P48168;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glycine receptor beta chain precursor (Glycine receptor 58 kDa  
 DE subunit).  
 GN GLRB.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95004576; PubMed=7920630;  
 RA Kingsmore S.F., Giros B., Suh D., Bieniarz M., Caron M.G.,  
 RA Seidlin M.F.;  
 RT "Glycine receptor beta-subunit gene mutation in spastic mouse  
 RT associated with LINE-1 element insertion.";  
 RL Nat. Genet. 7:136-141(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A. AND VARIANTS SPASTIC.  
 RC STRAIN=C57BL/6, and BALB/c; TISSUE=Brain, and Liver;  
 RX MEDLINE=9503198; PubMed=7946325;  
 RA Muehlhardt C., Fischer M., Gass P., Simon-Chazottes D.,  
 RA Guenet J.-L., Kuhse J., Betz H., Becker C.M.;  
 RT "The spastic mouse: aberrant splicing of glycine receptor beta  
 RT subunit mRNA caused by intronic insertion of L1 element.";  
 RL Neuron 13:1003-1015(1994).  
 CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION  
 CC CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE  
 CC CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF  
 CC NEURONAL FIRING).  
 CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA AND BETA SUBUNITS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

Query Match 69.6%; Score 1568.5; DB 1; Length 457;  
 Best Local Similarity 71.5%; Pred. No. 4.2e-126;

```
CC -1- TISSUE SPECIFICITY: HIGH LEVELS OF EXPRESSION IN CORTEX.
CC HIPPOCAMPUS, THALAMUS AND CEREBELLUM.
CC -1- DISEASE: DEFECTS IN GLRB CAUSE THE SPASTIC CONDITION WHICH IS
CC CHARACTERIZED BY MUSCLE RIGIDITY, TREMORS, MYOCLONIC JERKS,
CC PRONOUNCED STARTLE REACTION, ABNORMAL GAIT AND IMPAIRED RIGHTING
CC ABILITY.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```

```
DB 282 VLWSLSPWINDPASAARVPGLIFSVLSLASECTTLAELPKVSYVKALDWLLIACLLFGF 341
OY 345 ALLEYTAIINFVSRKHKEFRLRRORR 372
DB 342 ASLVEAYVQVMNLNPK---RVEAEKRR 366

RESULT 12
GRB_RAT
ID GRB_RAT STANDARD; PRT; 496 AA.
AC P20781;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glycine receptor beta chain precursor (Glycine receptor 58 kDa subunit).
DE subunit).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP MEDLINE=90297968; Pubmed=2163264;
RX Greeningloh G., Pribilla I., Prior P., Multaup G., Beyreuther K., Taleb O., Betz H.;
RA "Cloning and expression of the 58 kd beta subunit of the inhibitory RT glycine receptor."
RL Neuron 4:963-970(1990).
CC -1- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THIS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).
CC -1- SUBUNIT: PENTAMER COMPOSED OF ALPHA (48 kDa) AND BETA (58 kDa) SUBUNITS. A THIRD TYPE OF SUBUNIT (93 kDa) IS A PERIPHERAL MEMBRANE PROTEIN ASSOCIATED WITH THE CYTOPLASMIC DOMAINS OF THE RECEPTOR CORE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR PIR: JH0165; JH0165.
DR InterPro: IPR000188; GABA_A_receptor.
DR InterPro: IPR001175; Neur_chan.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PR00252; NRIONCHANNEL.
DR TIGRPFAMS: TIGR00860; LTC; 1.
DR PROSITE: PS00236; NEUROTR_ION_CHANNEL; 1.
DR Receptor: Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane.
FT SIGNAL 1 22
FT CHAIN 23 496
FT DOMAIN 23 265
FT TRANSEM 266 290
FT TRANSEM 299 316
FT TRANSEM 331 354
FT DOMAIN 355 477
FT TRANSEM 478 495
FT DISULFID 183 197
FT CARBOHYD 54 54
FT VARIANT 242 242
FT SEQUENCE 496 AA; 55927 MW; F77CE96105EF8ACF CRC64;

Query Match 38.3%; Score 863.5; DB 1; Length 496;
Best Local Similarity 47.8%; Pred. No. 5,7e-66;
Matches 187; Conservative 63; Mismatches 88; Indels 53; Gaps 11;

OY 9 LSFLLMTLPQGVLLRYALAKEEYKSGTSGSO-----PMSPSDLDK 50
DB 7 VSEFILMSL-----LFEDACSKESKSKGKKKQYLCPSQSOASDLARVPNPSTSLNLR 62
OY 51 LMGRTSGYDARIRNFKGPPVNVTCNIFINSFSSVTTMDYRVNVFLRQOQNDPRL--- 107
```

Db 63 LL---VSYDPRIRPNFKGIPVDVYVNFINSFGSIQETTMIDRYNIFLRQKNDPRLKLP 119  
 QY 108 -SYREYPPDSDLDPPSMIDSIWKPDLEFANEKANFHEVTITDNKLLRIFKNGVLYSIRL 166  
 Db 120 SDFR--GSDALVDPPTWKCKLMKPDLEFANEKANFHEVTIOENILLFFRGGDVLVSMRL 177  
 QY 167 TLILSCMLDKNPFMDIQTCTMQLSSSILCSPLSLSVGYTKMDLVFEMLEDAVAVQ 226  
 Db 178 SITLSCPLDLTLFPMDTQRCKMQLE-----SFGYTTDDLRFIMOSGDP-VQ 222  
 QY 227 VAEGTLTLPF--ILRDEKDLGCGCTKHV-NTGKFTCIEVKFHLERQMGYLLQMYIPSLIIV 284  
 Db 223 L-EKIALPQFDIKEDIEYGNCTKYKGTGYTCVEVIFTLRQVGYFMGVYAPPTLIV 281  
 QY 285 ILSWSEFMINMDAPARVGLGITTTLVMTTOSGSRASLPKVSYSYKADIMAVCLLFEV 344  
 Db 282 VLSWSEFMINPDASARVPLGIFSVLSLASECTTLAELPKVSYSYKALDVLMLACLLFGF 341  
 QY 345 AALLEYAAI----NFVSRQKKEFIRLRRROR 371  
 Db 342 ASLVEYAVVQVLMNPKRVEAEKARIKAEQ 372

## RESULT 13

GRB\_HUMAN STANDARD: PRT: 497 AA.

ID GRB\_HUMAN STANDARD: PRT: 497 AA.  
 AC P48167;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glycine receptor beta chain precursor (Glycine receptor 58 kDa subunit).  
 GN GLRB.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hippocampus;  
 RX MEDLINE=96352561; PubMed=8717357;  
 RA Handford C.A., Lynch J.W., Baker E., Webb G.C., Ford J.H., Sutherland G.R., Schofield P.R.;  
 RT "The human glycine receptor beta subunit: primary structure, functional characterisation and chromosomal localisation of the human and murine genes.";  
 RT Brain Res. Mol. Brain Res. 35:211-219(1996).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96341117; PubMed=9676428;  
 RA Milani N., Meilhardt C., Weber R.G., Lichter P., Kioschis P., Pouscka A., Becker C.-M.;  
 RT "The human glycine receptor beta subunit gene (GLRB): structure, refined chromosomal localization, and population polymorphism.";  
 RL Genomics 50:341-345(1998).  
 CC -I- FUNCTION: THE GLYCINE RECEPTOR IS A NEUROTRANSMITTER-GATED ION CHANNEL. BINDING OF GLYCINE TO ITS RECEPTOR INCREASES THE CHLORIDE CONDUCTANCE AND THUS PRODUCES HYPERPOLARIZATION (INHIBITION OF NEURONAL FIRING).  
 CC -I- SUBUNIT: PENTAMER COMPOSED OF ALPHA AND BETA SUBUNITS.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -I- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: U33267; AAB37750.1;  
 CC EMBL: AF094754; AAC71033.1;

DR EMBL: AF094755; AAC71034.1; -  
 DR Genew; HGNC:4329; GLRB.  
 DR MIM; 138492; -  
 DR InterPro: IPR000188; GABA\_receptor.  
 DR InterPro: IPR01175; Neur\_channel.  
 DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
 DR Pfam: PF02932; Neur\_chan\_membr; 1.  
 DR PRINTS: PR00252; NRIONCHANNEL.  
 DR TIGRFAMs: TIGR00860; LIC; 1.  
 DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal; Transmembrane.  
 KM SIGNAL 1 22  
 FT CHAIN 23 497  
 FT DOMAIN 23 265  
 FT TRANSMEM 266 290  
 FT TRANSMEM 299 316  
 FT TRANSMEM 331 354  
 FT DOMAIN 355 478  
 FT TRANSMEM 479 496  
 FT DISULFID 183 197  
 FT CARBOHYD 54 54  
 FT CARBOHYD 242 242  
 SQ SEQUENCE 497 AA; 56122 MW; C1F0B407601D3625 CRC64;

Query Match 38.0%; Score 855.5; DB 1; Length 497;  
 Best Local Similarity 47.4%; Pred. No. 2.7e-65;  
 Matches 186; Conservative 64; Mismatches 89; Indels 53; Gaps 11;

QY 8 TLSPILLTLTGLGQVLLRALAKEVSKSGTKSQP-----SPSDFLD 49  
 Db 6 TTAFLITL-----WVEEAYSKEKSSKKKGKQYOLCPSSQASADLAVPANSSTNLN 61  
 QY 50 KIMGRTSGYDARIRPNFKGPPVNVTCNIFISFSSVTKTMDRYNVNFLRQKNDPRL-- 107  
 Db 62 RLL---VSYDPRIRPNFKGIPVDVYVNFINSFGSIQETTMIDRYNIFLRQKNDPRLK 118  
 QY 108 -SYREYPPDSDLDPPSMIDSIWKPDLEFANEKANFHEVTITDNKLLRIFKNGVLYSIR 165  
 Db 119 PSDFR--GSDALVDPPTWKCKLMKPDLEFANEKANFHEVTIOENILLFFRGGDVLVSMR 176  
 QY 166 LTLILSCMLDKNPFMDIQTCTMQLSSSILCSPLSLSVGYTKMDLVFEMLEDAVAV 225  
 Db 177 LSTLSCPLDLTLFPMDTQRCKMQLE-----SFGYTTDDLRFIMOSGDP-V 221  
 QY 226 VAEGTLTLPF--ILRDEKDLGCGCTKHV-NTGKFTCIEVKFHLERQMGYLLQMYIPSLI 283  
 Db 222 QL-EKIALPQFDIKEDIEYGNCTKYKGTGYTCVEVIFTLRQVGYFMGVYAPPTLIV 280  
 QY 284 VLSWSEFMINMDAPARVGLGITTTLVMTTOSGSRASLPKVSYSYKADIMAVCLLFEV 343  
 Db 281 VLSWSEFMINPDASARVPLGIFSVLSLASECTTLAELPKVSYSYKALDVLMLACLLFG 340  
 QY 344 AALLEYAAI----NFVSRQKKEFIRLRRROR 371  
 Db 341 ASLVEYAVVQVLMNPKRVEAEKARIKAEQ 372

## RESULT 14

GAB\_LYMST STANDARD: PRT: 499 AA.

ID GAB\_LYMST STANDARD: PRT: 499 AA.  
 AC P26714;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Gamma-aminobutyric-acid receptor beta subunit precursor (GABA(A) receptor).  
 OS Lymnaea stagnalis (Great pond snail).  
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;  
 CC Lymnaeidae; Lymnaea.  
 OX NCBI\_TaxId=6523;  
 RN [1]  
 RP SEQUENCE FROM N.A.



FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 160 174 BY SIMILARITY.  
FT VARSPLIC 361 364 MISSING (IN ISOFORM BETA-4).  
SQ SEQUENCE 488 AA; 56068 MW; 164A75314BDB2C12 CRC64;

Query Match 32.2%; Score 726; DB 1; Length 488;

Best local similarity 38.0%; Pred. No. 2.9e-54;

Matches 156; Conservative 79; Mismatches 124; Indels 52; Gaps 10;

QY 43 SPS-----DELKIMGRTSGYDARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRV 94  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
DB 27 SPSTGNISVYKEIVDKL---KGYDRLRPDPGNPVTVMGMSIHSSIDQISEVNDYTI 83  
95 NVFLRQQWMDPRLSYREYPPDSLDLPPSLMSIMKPDLEFANEKGANFHEVTTDNKLRI 154  
84 TMYFOQSMRDKRLAYNDLP-LNLTLDNRYADQLMLPDYFELNDKKSFLHGVTVKNMIRL 142  
QY 155 FKNGVLYSIRLTLILSLMDLKNFPMIQTCTMOLESSIICSLPLSLSLSGYTMKD 214  
DB 143 HPDGTLYGLRITTTAAACMDLRRYPPLDQONCTLEIE-----SYGYTVDDI 188  
QY 215 VFEMLEDAVAVAGLTLPOFILRDEKDLGCTKHVNTGKPTCTEVKFLEROMGYLI 274  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
DB 189 VFFMOGNSAVTGMVLELPFTITEQR-LVSRVYFTTGSYLRSLSPRIKRNIGYFTL 247  
QY 275 QMYIPSLILVILSWSFWINMDAPARVGLITTVLMTTQSSGRASLPKYSYKAIDI 334  
||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :  
DB 248 QTYMPSILITLISWVSWINMDASARVALGTYTLMTTINTHLRETLPIPYKAIDV 307  
QY 335 KMAVCLLFFAALLLEYAINEV-----SRQKEFIRL-----RRQRORLEEDIOE 382  
DB 308 YLMGCFVFFELALLLEYAFVNYIFGSGRPOQKOSERISKANNERHRYEKEKRVREQVDPY 367  
QY 383 SRFYFRGYGLGHCLQARD-----GG-----PMESGIGYSPQAPLREG 422  
DB 368 GNILLSTLDNNNELLATDMSSVGDGRNSVMSFEGSGIOFRKPLAS--RDG 416

Search completed: June 25, 2003, 17:15:56  
Job time : 19.3557 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 25, 2003, 17:13:30 ; Search time 52.0671 Seconds  
(without alignments)  
1705.612 Million cell updates/sec

Title: US-10-075-846-4  
Perfect score: 2252  
Sequence: 1 MTLVPAFLSFLMLTLPQ.....PQPPAPLRGETTRKLYVD 431

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_plant:\*  
10: sp\_protist:\*  
11: sp\_virus:\*  
12: sp\_virus:\*  
13: sp\_unclassified:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1886	83.7	416	11	08VHF3 mus musculus
2	1749.5	77.7	456	13	08VHF3 mus musculus
3	1598	71.0	480	11	09J9C9 ratius norv
4	1578	70.1	463	11	091XP5 mus musculus
5	1570	69.7	449	6	09GKF0 mus musculus
6	1570	69.7	449	6	09GKF0 mus musculus
7	1555.5	69.1	459	13	08OC93 mus musculus
8	1537.5	68.3	450	13	09OW14 mus musculus
9	1381	61.3	426	13	09OW14 mus musculus
10	876.5	38.9	498	13	09OW14 mus musculus
11	850.5	37.8	497	6	09GJ59 mus musculus
12	747.5	33.2	486	5	09GJ59 mus musculus
13	743.5	33.0	438	5	046124 haemochus
14	742.5	33.0	453	5	077295 haemochus
15	730.5	32.4	456	5	094900 drosophila
16	729.5	32.4	452	5	076773 lucilia cup

17	729	32.4	606	5	09VSV0 drosophila
18	728	32.3	606	5	09BLX8 drosophila
19	724.5	32.2	481	5	018471 heliothis v
20	724.5	32.2	496	5	018468 heliothis v
21	724	32.1	533	5	016896 aedes aegypt
22	722.5	32.1	456	5	09G053 drosophila
23	722	32.0	430	5	09G097 caenorhabdi
24	721	32.0	430	5	017369 caenorhabdi
25	720	32.0	601	5	092138 drosophila
26	719.5	31.9	454	5	09VDS5 drosophila
27	718	31.9	421	5	046123 haemochus
28	717	31.8	454	5	0961R4 drosophila
29	716	31.8	430	13	090590 gallus gall
30	713	31.7	541	5	09U9B8 lucilia cup
31	712.5	31.6	550	5	017145 lucilia cup
32	711.5	31.6	494	5	09X2W0 heliothis v
33	710.5	31.5	416	5	09Y096 caenorhabdi
34	708.5	31.5	416	5	017367 caenorhabdi
35	708	31.4	336	5	09G052 drosophila
36	706.5	31.4	351	5	025634 onchocerca
37	705.5	31.3	1106	5	022637 caenorhabdi
38	704	31.3	379	5	096964 ascaris suu
39	700.5	31.1	478	5	095P46 caenorhabdi
40	700.5	31.1	657	5	09TW41 caenorhabdi
41	700	31.1	519	13	09DD99 brachydanio
42	695.5	30.9	478	5	017548 caenorhabdi
43	695.5	30.9	657	5	017547 caenorhabdi
44	690.5	30.7	435	5	09TY66 haemochus
45	690	30.6	484	5	09U990 caenorhabdi

## ALIGNMENTS

RESULT 1  
ID 08VHF3 PRELIMINARY: PRT: 416 AA.  
AC 08VHF3:  
DT 01-MAR-2002 (TREMBLrel. 20. Created)  
DT 01-MAR-2002 (TREMBLrel. 20. Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21. Last annotation update)  
DE Glycine receptor alpha 4 subunit.  
GN GLRA4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BC3/FE; TISSUE=SPINAL CORD;  
RA Groemer T.-W.W., Becker C.-M.M., Becker K.;  
RT "Localization of different glycine receptor isoforms in murine spinal cord."  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF462147; AAL69899.1; -  
DR InterPro: IPR001187; GABA\_A\_receptor.  
DR InterPro: IPR001175; Neur\_chan.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR PRINTS: PF02932; Neur\_chan\_membr; 1.  
DR TIGRFAMs: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; UNKNOWN\_1.  
KW Receptor.  
SO SEQUENCE 416 AA; 48322 MW; C095998PAC077451 CRC64;

Query Match 83.7%; Score 1886; DB 11; Length 416;  
Best Local Similarity 91.8%; Pred. No. 3.7e-169;  
Matches 358; Conservative 10; Mismatches 8; Indels 14; Gaps 1;

QY 42 MSPSFLDKMGKGRISGVYARIRPNKGPVAVNTCHIFINSSSYTKTMDYRVNVPFLQO 101  
DB 1 MSPSFLDKMGKGRISGVYARIRPNKGPVAVNTCHIFINSSSYTKTMDYRVNVPFLQO 60



```
Db 144 KPDLFANEKGANHEVTTDKLIRFKNGNVLYSIRLTLLSCPMDLKNPMDVQTCIM 203
Oy 189 QLESSLILCSPLSPSLSTVGTMDLVFEMLEDAPAVOVAGLTLPOFLRDEKDLGCT 248
Db 204 QLE-----SFGYTMNDLIFEWODEAP-VQVAGLTLPOFLRDEKDLGCT 248
Oy 249 KHYTGKFTCIYEVFHLEROMGYLLIOMYIPSLILVILSVSPFINDDAPARALGITT 308
Db 249 KHYTGKFTCIYEVFHLEROMGYLLIOMYIPSLILVILSVSPFINDDAPARALGITT 308
Oy 309 VLTMTTOSGSRASLPKVSYYKAIIDIMWAVCLLFVFAALLEYAAINFVSRQHKFFIRLR 368
Db 309 VLTMTTOSGSRASLPKVSYYKAIIDIMWAVCLLFVFAALLEYAAINFVSRQHKFFIRLR 368
Oy 369 RQRQ-----RLE-----EDIIQESRFRFGYGLGHCLQARG-----GPMESGITYSP 412
Db 369 KRKKKTEAFALFKYRFSDDEVRFSRFTAYGMGPCLOAKGVVPPKGNHAYV--- 425
Oy 413 QPPAPLREGETTRKLYVD 431
Db 426 -----MKRSADMKKVFID 439
```

## RESULT 4

```
ID 091XP5 PRELIMINARY; PRT: 463 AA.
AC 091XP5;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycine receptor alpha 3 subunit...
GN GLRA3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SPINAL CORD;
RA Noegel S., Becker C., Becker K.;
RT "Different glycine receptor isoforms are expressed in murine
RT cerebellum."
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF362764; AAK51962.1; -.
DR MGD: MGI:95749; Glra3.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_membr; 1.
DR PRINTS: PRO1620; GABAARGAMMA.
DR TIGRFAMs: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEUROR_LON_CHANNEL; UNKNOWN.1.
KW Receptor.
SQ SEQUENCE 463 AA: 53605 MW: 889D67DE9B223605 CRC64;
```

Query Match 70.1%; Score 1578; DB 11; Length 463;  
Best Local Similarity 73.0%; Pred. No. 4.5e-140;  
Matches 311; Conservative 33; Mismatches 42; Indels 40; Gaps 8;

```
Oy 22 LLRLAALKEEYKSGTKSGQSPSPDLDKLGRTSGYDARIRPNFKGPPVAVTNCIFINS 81
Db 21 LLSLVATKETSARSRSAPSPDLDKLGRTSGYDARIRPNFKGPPVAVTNCIFINS 80
Oy 82 FSSVTKTTMDYRVAVFLRQONNDRLSTREYPPDSDLDPMSLDSIKRPDLFFANEKAN 141
Db 81 FGSIAETTMADYRVAVFLRQONNDRLSTREYPPDSDLDPMSLDSIKRPDLFFANEKAN 139
Oy 142 FHEVTTDKRLIRFNGNVLYSIRLTLLSCPMDLKNPMDVQTCIMLESSSILCSPL 201
Db 140 FHEVTTDKRLIRFNGNVLYSIRLTLLSCPMDLKNPMDVQTCIMLE----- 189
Oy 202 SLISVGYTKMDLVFEMLEDAPAVOVAGLTLPOFLRDEKDLGCTKHYTGKFTCIY 261
```

```
Db 190 ----SFGYTMNDLIFEWODEAP-VQVAGLTLPOFLRDEKDLRYCTKHYTGKFTCIY 244
Oy 262 KFHLEKRGYLLIOMYIPSLILVILSVSPFINDDAPARALGITTTLTMTTOSSSSRA 321
Db 245 RFLEROMGYLLIOMYIPSLILVILSVSPFINDDAPARALGITDMLTMTTOSSSSRA 304
Oy 322 SLKVSYYKAIIDIMWAVCLLFVFAALLEYAAINFVSRQHKFFIRRRORRO-----RLE- 376
Db 305 SLKVSYYKAIIDIMWAVCLLFVFAALLEYAAINFVSRQHKFFIRRRORRO-----RLE- 376
Oy 377 -----EDIIQESRFRFGYGLGHCLQARG-----GPMESGITYSPQPPAPLREGETT 425
Db 365 FYRSDPDDEVRFSRFTAYGMGPCLOAKGVVPPKGNHAYV---MKSP-----DEM 416
Oy 426 RKLKLYVD 431
Db 417 RKVFID 422
```

## RESULT 5

```
ID 09GKFO PRELIMINARY; PRT: 449 AA.
AC 09GKFO;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Glycine receptor alpha 1 subunit Isoform b.
GN GLRA1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21109390; PubMed=11178872;
RA Pierce K.D., Handford C.A., Morris R., Vafa B., Dennis J.A.,
RA Healy P.J., Schofield P.R.;
RT "A nonsense mutation in the alpha subunit of the inhibitory glycine
RT receptor associated with bovine myoclonus."
RL Mol. Cell. Neurosci. 17:354-363(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL: AF268366; AAG41140.1; -.
DR EMBL: AF268358; AAG41140.1; JOINED.
DR EMBL: AF268359; AAG41140.1; JOINED.
DR EMBL: AF268360; AAG41140.1; JOINED.
DR EMBL: AF268361; AAG41140.1; JOINED.
DR EMBL: AF268362; AAG41140.1; JOINED.
DR EMBL: AF268363; AAG41140.1; JOINED.
DR EMBL: AF268364; AAG41140.1; JOINED.
DR EMBL: AF268365; AAG41140.1; JOINED.
DR InterPro: IPR000188; GABAA_receptor.
DR InterPro: IPR001175; Neur_channel.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_membr; 1.
DR PRINTS: PRO1620; GABAARGAMMA.
DR PRINTS: PRO0252; NRIONCHANNEL.
DR TIGRFAMs: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEUROR_LON_CHANNEL; 1.
KW Glycopolysaccharide; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 449 AA: 51676 MW: DC75FB56F3114A4C CRC64;
```

Query Match 69.7%; Score 1570; DB 6; Length 449;  
Best Local Similarity 72.6%; Pred. No. 2.5e-139;  
Matches 307; Conservative 37; Mismatches 53; Indels 26; Gaps 8;

```
Oy 12 LLMTLTPROQVLLRLAALKEEYKSGTKSGQSPSPDLDKLGRTSGYDARIRPNFKGPPV 71
Db 9 LYLWE--TIVFSLAASKAEAAARSASKPSPDLDKLGRTSGYDARIRPNFKGPPV 65
```

```
QY 72 NTCNIFINSFSSVTKTMDYRVNVEFLRQWMDPRLSYREYDSDLDLDPMSLDSIMKPD 131
      |||:|||||:|:|||||:|||||:|:|||||:|||||:|:|||||:|||||:|:|||||:
Db 66 NVSCNIFINSFSGSIAETTDYRVNVEFLRQWMDPRLXAYNEYPDDSLDLPMSLDSIMKPD 125
QY 132 LFFANEKGANFHEVTTDNKLLRIFKNGVNLVIRLTLLISCLMDLKNFPMQOTQTMOLE 191
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 126 LFFANEKGANFHEITTDNKLXISNGVNLVIRLTLLACPMDLKNFPMQOTQTMOLE 185
QY 192 SSSIIICSPLSLSVGYTMDLVEFWEMLEDAVAVAGELTLPQFLRDEKDLGCTKH 251
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 166 -----SFGYTMNDLIFEWQEOG-AVQVADLTLPQFLKEKDLRYCTKH 230
QY 252 NTGKFTCTEVRKFLEROMGYLLIOMYIPSLIVILSWSEWIMMDAAPARVGLGTTVLT 311
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 231 NTGKFTCTEVRKFLEROMGYLLIOMYIPSLIVILSWSEWIMMDAAPARVGLGTTVLT 290
QY 312 MTTQSSGSRASLPKYSYKVIDIMMAVCLLFEVALLLEVAALNPFVSRQKEIRLROR 371
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 291 MTTQSSGSRASLPKYSYKVIDIMMAVCLLFEVALLLEVAALNPFVSRQKEIRLROR 350
QY 372 RORLEEDIIQESRFYFRGYGLG-HCLQARDGGMGEG---SGIYSPQAPAPLREGETTRK 427
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 351 HHK--EDEAGEGRFNFSAVGMGPACLOAKDGISVKGANNSNTNP-PPAP-SKSEPEMK 406
QY 428 LKY 430
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 407 LFI 409
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:

RESULT 6
Q9GKE9 PRELIMINARY: PRT: 457 AA.
ID Q9GKE9
AC Q9GKE9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Glycine receptor alpha 1 subunit isoform a.
GN GLRA1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21109390; PubMed-11178872;
RA Pierce K.D., Handford C.A., Morris R., Vafa B., Dennis J.A.,
RA Healy P.J., Schofield P.R.,
RT "A nonsense mutation in the alpha subunit of the inhibitory glycine
RT receptor associated with bovine myoclonus.";
RT Mol. Cell. Neurosci. 17:354-363(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL: AF268366; AAG41141.1; -.
DR EMBL: AF268358; AAG41141.1; JOINED.
DR EMBL: AF268359; AAG41141.1; JOINED.
DR EMBL: AF268360; AAG41141.1; JOINED.
DR EMBL: AF268361; AAG41141.1; JOINED.
DR EMBL: AF268362; AAG41141.1; JOINED.
DR EMBL: AF268363; AAG41141.1; JOINED.
DR EMBL: AF268364; AAG41141.1; JOINED.
DR EMBL: AF268365; AAG41141.1; JOINED.
DR EMBL: AF268366; AAG41141.1; JOINED.
DR InterPro: IPR000188; GABA_A_receptor.
DR InterPro: IPR001175; Neur_chan.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PR01620; GABAR_GAMMA.
DR PRINTS: PR00252; NRIONCHANNEL.
DR TIGRFS: TIGR00860; LIG_1.
DR PROSITE: PS00236; NEURON_ION_CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 457 AA: 52607 MW: F3IC41BFCS786908 CRC64;
```

```
Query Match 69.7%; Score 1570; DB 6; Length 457;
Best Local Similarity 71.6%; Pred. No. 5.9e-138;
Matches 307; Conservative 38; Mismatches 54; Indels 30; Gaps 8;

QY 12 LLLMTLPQGVLLRVALAKEEVKSGTKGSPSPDFLKMGRISGYDARIRPNP----- 66
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 9 LYLWE---TIVFFSLAASKAEAAARSASPKSPDFLKMGRISGYDARIRPNKGPV 65
QY 72 NTCNIFINSFSSVTKTMDYRVNVEFLRQWMDPRLSYREYDSDLDLDPMSLDSIMKPD 131
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 66 NVSCNIFINSFSGSIAETTDYRVNVEFLRQWMDPRLXAYNEYPDDSLDLPMSLDSIMKPD 125
QY 132 LFFANEKGANFHEVTTDNKLLRIFKNGVNLVIRLTLLISCLMDLKNFPMQOTQTMOLE 191
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 126 LFFANEKGANFHEITTDNKLXISNGVNLVIRLTLLACPMDLKNFPMQOTQTMOLE 185
QY 192 SSSIIICSPLSLSVGYTMDLVEFWEMLEDAVAVAGELTLPQFLRDEKDLGCTKH 251
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 166 -----SFGYTMNDLIFEWQEOG-AVQVADLTLPQFLKEKDLRYCTKH 230
QY 252 NTGKFTCTEVRKFLEROMGYLLIOMYIPSLIVILSWSEWIMMDAAPARVGLGTTVLT 311
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 231 NTGKFTCTEVRKFLEROMGYLLIOMYIPSLIVILSWSEWIMMDAAPARVGLGTTVLT 290
QY 312 MTTQSSGSRASLPKYSYKVIDIMMAVCLLFEVALLLEVAALNPFVSRQKEIRLROR 371
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 291 MTTQSSGSRASLPKYSYKVIDIMMAVCLLFEVALLLEVAALNPFVSRQKEIRLROR 350
QY 372 ROR-----LEEDIIQESRFYFRGYGLG-HCLQARDGGMGEG---SGIYSPQAPAPLRE 421
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 351 HHKSPMLNLFQDEDEGEGRFNFSAVGMGPACLOAKDGISVKGANNSNTNP-PPAP-SKS 408
QY 422 GETTRKLY 430
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 409 PEMRKLF 417
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:

RESULT 7
Q9G93 PRELIMINARY: PRT: 459 AA.
ID Q9G93
AC Q9G93;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Glycine receptor alpha2L subunit.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Devignot V., Prado de Carvalho L., Bregestovski P., Goblet C.;
RT "A Novel Glycine Receptor  $\alpha$ 2L Subunit Variant in the Zebrafish
RT Brain.";
RT Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF488379; AAM00910.1; -.
KW Receptor.
SQ SEQUENCE 459 AA: 52330 MW: FFC97E289FA3404A CRC64;

Query Match 69.1%; Score 1555.5; DB 13; Length 459;
Best Local Similarity 69.4%; Pred. No. 5.9e-138;
Matches 302; Conservative 43; Mismatches 53; Indels 37; Gaps 6;

QY 12 LLLMTLPQGVLLRVALAKEEVKSGTKGSPSPDFLKMGRISGYDARIRPNP----- 66
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 6 LYLWE---TIVFFSLAASKAEAAARSASPKSPDFLKMGRISGYDARIRPNKGPV 61
QY 67 -----KQPPNVATCNIFINSFSSVTKTMDYRVNVEFLRQWMDPRLSYREYDSD 116
      |||:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:|:|||||:
Db 62 QGSKSDGTHKKGPVNVATCNIFINSFSGSIAETTDYRVNVEFLRQWMDPRLSYREYDSD 121
```

```

Oy      11  LDDPMDLDSINKPDLFEFNEKGFHEVTTDNKLLRFKGNVLYISRLTLLISCMDL 176
Db      122  LDDPMDLDSINKPDLFEFNEKGFHEVTTDNKLLRSKGNVLYISRTILVYACMDL 181

Oy      177  KNFMDIOTCTMOLESSLILCSPLPSLSVSGYTKMDLVEFWLEDAPAVOYAGLTLPOF 236
Db      182  KNFMDVOYCTIMOLE-----SEGYTNDYIDIFEM-DEKGAVOYAGLTLPOF 226

Oy      237  ILDEKDLCCCTKHYWTGKFTCIIEVKPHLEROMGYLLQWYIPSLLYIISWVSFWINMD 296
Db      227  ILKEEKDLVYCTKHNTGKFTCIEARPHLEROMGYLLQWYIPSLLYIISWVSFWINMD 286

Oy      297  AAPARGLCTITVLTMTTSSGSRSLSRKSVYKAIIDIMAVCLLFVPAALLEAYAINFV 356
Db      287  AAPARGLCTITVLTMTTSSGSRSLSRKSVYKAIIDIMAVCLLFVSALEEFVAAVNFI 346

Oy      357  SROCKEFILRRRORLEEDIIOESFRFYFGYGLG-HCIQARDGPMESGICVPOF 415
Db      347  ARQHKELLRROR--RRRHLEKDEADGRSFYAAYGMPACIQARDGMAIKGNNNNAPTST 404

Oy      416  APLLRGCTTRKLYV 430
Db      405  NPPEKTVEEKRLKI 419

```

	RESULT	8
DB	090W14	
ID	090W14	PRELIMINARY; PRT; 450 AA.
AC	090W14:	
DT	01-DEC-2001 (TrEMBLrel). 19,	Created)
DT	01-DEC-2001 (TrEMBLrel).	19, last sequence update)
DT	01-JUN-2002 (TrEMBLrel).	21, last annotation update)
DE	Glycine receptor alpha23 subunit precursor.	
GN	GLRA3 OR GLYRALPHA3.	
OS	Baccharodonta rerio (zebrafish) (zebra danio).	
OC	Eukaryota; Metazoa; Chordata; Craniata;	Vertebrates; Euteleostomi;
OC	Achnopterygii; Neopterygii; Teleostei;	Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.	
OX	NCBI_TaxID=7955.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TJSSUP-BRAIN:	
KX	MEDLINE=21541038; PubMed=11685575;	
RA	Imboden M., Devignot V., Goblet C.;	
RT	"Phylogenetic relationships and chromosomal location of five distinct	
RT	glycine receptor subunit genes in the teleost danio rerio.";	
RL	Dev. Genes Evol. 211:415-422(2001).	
DR	EMBL: AJ308516; CACC38837.1; -	
DR	ZFIN: ZDB-GENE-020402-1; glra3.	
DR	InterPro: IPR000188; GABA_A_receptor.	
DR	InterPro: IPR001175; Neur_channel.	
DR	Pfam: PF02931; Neur_chan_LBD; 1.	
DR	Pfam: PF02932; Neur_chan_memb; 1.	
DR	PRINTS: PR01620; GABARGAMMA.	
DR	TIGRFAMS: TIGR00860; LIC; 1.	
DR	PROSITE: PS00236; NEUROTR_ION_CHANNEL; UNKNOWN_1.	
KM	Receptor; Signal.	
FT	SIGNAL	1..27 POTENTIAL.
FT	CHAIN	28..450 GLYCINE RECEPTOR ALPHA23 SUBUNIT.
ST	SEQUENCE	450 AA; 51636 MW; FDD64C09FEF60EAO CRC64;
DQ	Query Match	68.3%; Score 1537.5; DB 13; Length 450;
DQ	Best Local Similarity	72.9%; Pred. No. 2.9e-136;
DQ	Matches 301; Conservative 33;	Mismatches 54; Indels 25; Gaps 5
DQ	22 LNRVALAEKSVSGTGSQPMSPSDFLDKLGMRTSGYDARIRPNFKGPVNVTCNIPIINS	81
DQ	: :	81
DQ	21 ILTSIVASKEPESPERRAAPSPDPFLDKLMGRTSGYDARIIPNFKGPPVVNTCNIFINS	80
DQ	82 FSSATKTTHMYRYNNVLROOANDPRLSYEVEPDSDSLDLPSSLSTSINKRPDLFANEGKAN	144
DQ	I :	144
DQ	81 FGSIETHTMYRNITLRCKMDMPRIAYSEVPDSDSLDLPSSLSTSINKRPDLFANEGKAN	140

```

QY 142 FHEVTYDNKILRFKKGANVYSIRLTLLISCMDKPNPMDIOCTMDESSILCSPLP 201
DQ 141 FHEVTYDNKILRFKKGANVYSIRLTLLISCMDKPNPMDIOCTMDESSILCSPLP 190
QY 202 SUSLSVGYTMKDLVFEMLEBADAPOVAEBGLTLPQELTDERKDLCCCTKHNTGKFCIEV 261
DQ 191 ----SGCYTMNDLIFEMQCKGP--YOVAGGLTLPQELTDERKDLCCCTKHNTGKFCIEV 245
QY 262 KHFLEOMKGYLIOWYITPSLLVITLSWVSFWITMDAARAGLGTTLVMTPTQSSSRA 321
DQ 246 RHFLEOMKGYLIOWYITPSLLVITLSWVSFWITMDAARAGLGTTLVMTPTQSSSRT 305
QY 322 SLPRYSVAIAIDIMAAVCLLFVFAALLLEYAALINFSVOKHEFIRLRKOROROLEEDIIQ 381
DQ 306 SLPRYSVAIAIDIMAAVCLLFVFAALLEYAANVPSYOHKELTRRRRRRKSCKKEE--VR 364
QY 382 ESFRFYRGRGIGLGHCLDARDGCPMESSGITSPOPAPDLREG---ETTPKLYVD 431
DQ 365 ESFRFSTTPNA-----GKDGAVPKTANNAATTPSEPAVAPGSHDEMKKLTID 411

```

ID	090WT3	PRELIMINARY:	PRT:	426 AA.
AC	090WT3:			
DT	01-DEC-2001 (TREMBlrel. 19, Created)			
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)			
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)			
DE	Glycine receptor alpha4 subunit (Fragment).			
GN	GLRA4B OR GLYRALPHA4.			
OS	Brachydanio rerio (Zebrafish) (Zebra danio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
NC	NCBI_TaxID=7955;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BRAIN;			
RA	MEDLINE=21541038; PubMed=11685575;			
RT	Imboden M., Devignot V., Goblet C.;			
RT	"Phylogenetic relationships and chromosomal location of five distinct			
RL	glycine receptor subunit genes in the teleost danio rerio.";			
DR	Dev. Genes Evol. 211:415-422(2001).			
DR	EMBL; AJ308517; CAC38838.1;			
DR	ZFIN; ZDB-GENE-020402-2; glra4b.			
DR	InterPro; IPR000188; GABAA_receptor.			
DR	InterPro; IPR001175; Neur_channel.			
DR	Pfam; PF02931; Neur_chan_LBD; 1.			
DR	Pfam; PF02932; Neur_chan_membr; 1.			
DR	TIGRFAMS; TIGR00860; LIC; 1.			
DR	PROSITE; PS00236; NEUROTR_ION_CHANNEL; UNKNOWN_1.			
KW	Receptor.			
FT	NON_TER			
FT	SEQUENCE	426 AA;	48503 MW;	9187E20B873234B CRC64;
Query Match		61.3%;	Score 1381;	DB 13; Length 426;
Best Local Similarity		67.7%;	Pred. NO. 1.5e-121;	
Matches	273;	Conservative	31;	Mismatches 41; Indels 58; Gaps 5;
QY	68	GPANVNCNFINSSSVTKTMDYRVNVVPLRQONNDRLSVREVPDDSLDIDPSMLDSI	127	
DB	1	GPANVNCNFINSSSVTKTMDYRVNVVPLRQONNDRLSVREVPDDSLDIDPSMLDSI	60	
QY	128	WKPDLEFANEKGANHEVTTNKKLIRFKNGNVLSIRLTLLSCPLMDKKNPMDIOTCT	187	
DB	61	WKPDLEFANEKGANHEVTTNKKLIRFONGNVLSIRLTLLSCPLMDKKNPMDIOTCT	120	
QY	188	MOLESSSILCSPLPSLSLSVGYTMDKLFWEMLDPAVQVAGLTLPOFILDEKDLCC	247	
DB	121	VQLE-----SFGYTNMDLLEQWMLDEGP-VQVADMDMLPQFVLKEEKDLGYC	165	
QY	248	TKHNITGCTCLEVAFHLERQMGYLIQMYIPSLILVILSVSWFINNDAPARVGLGIT	307	

Db 166 ANHYTGKFTGCTIEVKFHLEROMAYLLIOMYIPSLTLLVLSWVSWFMINDAARVGLGIT 225  
QY 308 TVLTMTTSSGSRASLPKSYVKAIDIMAVCLLFVFAALLLEYAAINFSROHKEFRLR 367  
Db 226 TVLTMTTSSGSRASLPKSYVKAIDIMAVCLLFVFAALLLEYAAINFSROHKEFRLR 285  
QY 368 RRORO-----RLEDDIIQESFFPR 388  
Db 286 RRLROQRNRAASGQGAETKNSNNVTGTPSRNAORQCSACAREEQSLASQNDLPP 345  
QY 389 GYGLCHICLARDGGPMEGSGISYPOPPALLREGETRLTYD 431  
Db 346 GFGIDTSLSG-DGPLESAAMFAGLPPHALFD---IRREVE 384

## RESULT 10

Q9DES8 PRELIMINARY: PRT: 498 AA.

AC 09DES8: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
GN Glycine receptor beta2 subunit precursor (Fragment).  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_Taxid=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE-BRAIN:  
RA Imboden M., Devignot V., Korn H., Goblet C.;  
RT Regional Distribution of Glycine Receptor mRNA in the Central Nervous  
RT System of Zebrafish."  
Submitted (Jul-2000) to the EMBL/Genbank/DBI databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (By similarity).  
-1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
EMBL: AJ404971; CC61688.1;  
DR ZFIN: ZDB-GENE-010410-2; g1rb.  
DR InterPro: IPR000188; GABA\_A\_receptor.  
DR InterPro: IPR001175; Neur\_channel.  
DR Pfam: PF02932; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_Memb; 1.  
DR PRINTS: PRO1620; GABARGAMA.  
DR PRINTS: PRO0252; NRIONCHANNEL.  
DR TIGRFAMs: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEURORF\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor; Signal;  
KW Transmembrane.  
FT NON\_TER 1  
FT SIGNAL <1 22  
SQ SEQUENCE 498 AA: 56358 MW: 12436FC344F0BB CRC64;

Query Match 38.9%; Score 876.5; DB 13; Length 498;  
Best Local Similarity 49.1%; Pred. No. 6, 4e-74;

Matches 189; Conservative 59; Mismatches 96; Indels 41; Gaps 11;

QY 17 LFGQVLLR--VALAKEE--VKSQTGSGSPMSPSDF-----LDKIMRTSG-----Y 58  
Db 8 LKGLITLMLLVQFSAEGRKPKKGGKGVICPSQSLASADLDVRVPA NSTSNILNRLMTY 67  
QY 59 DARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRVNVELRQOOWNDPRLSY-REYPDLSL 117  
Db 68 DSRIRNFKGIEVEDKVINIFINSFGSIQETMDYRVNIFLRQWMDPRLRLPDRFSDAL 127  
QY 118 DLDPSLDSIKRDLFFANEGKANHEVTTDKLIRFKNGVNLISIRLTLSCLMDIK 177  
Db 128 TVDPKFKQCLMKRDLFFANEGKANHEVTTDKLIRFKNGVNLISIRLTLSCLMDIK 187  
QY 178 NFPMOTCTMOLESSSIICSPPLSLSLSGVTMDLVFEWLEDAVAVAGLTLPQRT 237  
Db 188 LFPMDTQYCKMOLE-----SFGITTKDLVFMWOSGDP-VOMDE-TALPQD 231

QY 238 LRDEK-DLGCCTKHY-NTGKFTGCTIEVKFHLEROMGYLLIOMYIPSLTLLVLSWVSPWIM 295  
Db 232 VQEDLRKANCRTKFTYGTCTYTCVEYIFTLRQVGYMGVAFPTLLIVLSLWSRWNP 291  
QY 296 DAAPAVGLGTTVLTMTTSSGSRASLPKSYVKAIDIMAVCLLFVFAALLLEYAAI-- 353  
Db 292 DASARVPLGLTSLVLSSECTSLASELPKSYVKAIDIMAVCLLGFASLVEYAVGV 351  
QY 354 --NFSROHKEFRLRROROLE 376  
Db 352 MUNSFKRIEAKIKMAKEKAREKE 376

## RESULT 11

Q9GJS9 PRELIMINARY: PRT: 497 AA.

AC 09GJS9: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
GN Glycine receptor beta subunit.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_Taxid=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pierce K.D., Handford C.A., Morris R., Vafa B., Dennis J.A.,  
RA Healy P.J., Schofield P.R.;  
RT "Bovine Myoclonus is Caused by a Mutation in the  $\alpha 1$  Subunit of the  
RT Inhibitory Glycine Receptor."  
Submitted (May-2000) to the EMBL/Genbank/DBI databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (By similarity).  
-1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
EMBL: AF268376; AAG14345.1;  
DR EMBL: AF268374; AAG14345.1;  
DR EMBL: AF268367; AAG14345.1; JOINED.  
DR EMBL: AF268368; AAG14345.1; JOINED.  
DR EMBL: AF268369; AAG14345.1; JOINED.  
DR EMBL: AF268370; AAG14345.1; JOINED.  
DR EMBL: AF268371; AAG14345.1; JOINED.  
DR EMBL: AF268372; AAG14345.1; JOINED.  
DR EMBL: AF268373; AAG14345.1; JOINED.  
DR InterPro: IPR000188; GABA\_A\_receptor.  
DR InterPro: IPR001175; Neur\_channel.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_Memb; 1.  
DR PRINTS: PRO1620; GABARGAMA.  
DR PRINTS: PRO0252; NRIONCHANNEL.  
DR TIGRFAMs: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEURORF\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;  
KW Transmembrane.  
SQ SEQUENCE 497 AA: 56039 MW: 82F140C115A87E6 CRC64;

Query Match 37.8%; Score 850.5; DB 6; Length 497;  
Best Local Similarity 47.1%; Pred. No. 1, 8e-71;

Matches 184; Conservative 65; Mismatches 89; Indels 53; Gaps 11;

QY 9 LSFLLMTLPGQVLLVALAKEEVSGTGSGSPM-----SPSPFLDK 50  
Db 7 VAFETLISL-----CLIEAIVSKSKSKGKKGVICPSQSLASADLDARVPA NSTSNILN 62  
QY 51 LKGRISGYDARIRPNFKGPPVNTCNIFINSFSSVTKTMDYRVNVELRQOOWNDPRL--- 107  
Db 63 LL--VSYPRIAPNFKGIPDVVNVNIFINSFGSIQETMDYRVNIFLRQWMDPRLKLP 119  
QY 108 -SYREYPDSDLDLPSMDSIKRDLFFANEGKANHEVTTDKLIRFKNGVNLISIRL 166  
Db 120 SDFR--GSDALTVDPTRYCKMKRDLFFANEGKANHEVTTDKLIRFKNGVNLISIRL 177

Oy	167	TLTSLCLMDLNPFPDIDOTCMOWLESSILCSPLPSLSLVGYTKMKDLVFEWLEDAPAVO	226
Dd	178	SITSLSCPDLTLFFPDTRCKMQLF-----SFGYTDLLRFTIWSGDP-VQ	222
Oy	227	VAEGTLTPQF-TLRDEKDLGGCTKHV-WTGKFTCIETVEFHIEROMGYLIOMYIPSLLTIV	284
Dd	223	L-EKIALPQFIKKEDIEYGNCTKYKGATGYTCVEVFETLARQGVFMWGYPAPLTLLIV	281
Oy	285	ILSVSEFINNDAPARAGLGITTVYLMTTOSSGSRASLPKYSYKAIDIMKAVCLLFVE	344
Dd	282	VLSMWISFENINPDASAARYPLGFVSLSLASRECTTLAAELPKYSYKALDWMLACLLEQF	341
Oy	345	AALLEYAII-----NFVSROHKEFIRLROR 371	
Dd	342	ASLVEYIAVVQYMNLNPKRYEAKRIARAEO 372	
 RESULT 12 O9GYU4 ID PRELIMINARY: PRT: 486 AA.			
AC	O9GYU4:		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	GABA-A receptor beta subunit.		
OC	Septia officinalis (Common cuttlefish).		
OC	Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Sepioidae;		
OC	Seplidae; Sepia.		
OX	NCBI_TaxID=6610;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kirby R.R., Williamson R., Farley S.;		
RT	"GABA-A receptor gene sequence from the central nervous system of the		
RT	cuttlefish, Sepia officinalis.";		
RL	Abstr. - Soc. Neurosci. 23:957-957(1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Kirby R.R.;		
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.		
DR	EMBL: AY005810; AAF97816.1; "		
DR	InterPro: IPR000188; GABAA_receptor.		
DR	InterPro: IPR001175; Neur_channel.		
DR	Pfam: PF02931; Neur_chan_LBD.		
DR	Pfam: PF02932; Neur_chan_membr_1.		
DR	PRINTS: PR01620; GABAARGAMMA.		
DR	PRINTS: PR00252; NRIONCHANNEL.		
DR	TIGRFAMS: TIGR00860; LIC; 1.		
DR	PROSITE: PS00236; NEURORION_CHANNEL. 1.		
KW	Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;		
KW	Transmembrane.		
SO	SEQUENCE 486 AA; 56640 MW; 1FD6DBCCE1C7257 CRC64;		
 Query Match 33.2%; Score 747.5; DB 5; Length 486; Best Local Similarity 42.4%; Pred No. 8.8e-62; Matches 148; Conservative 69; Mismatches 109; Indels 23; Gaps 5			
Oy	9	LSFLLLMTLPGCVLLRVALAKEEVKSGTKGSQP--MSPSDFLDKLNGRTSGYDARIKPNF	66
Dd	3	LKIYAMLVLKESTVLITLV---LLVGSYGDGDSRRONTDTIERLL--KGYDIRLRPQF	56
Oy	67	KGPPIVNVICNFINFSSTYTKTITMYRVNVLFRQONDPRLSYREYPDSDLSDLPESMLDS	126
Dd	57	EYQAARIGMEIIASFDSISEVMNYDYTTIMYNQYWTDERLYVSNDSNDNLTLTGDPFAEK	116
Oy	127	IWKPPDLPEANEGKANFHEVTYTNKDLIRFKNGCNVLSIFLTILTSLCDLMKNPMPDIQC	186
Dd	117	IWPVPPTFRANKDNNSLHDITTEKNKMTRLFGNSIYGMKFITTLACMDLHYFLPDEQNC	176
Oy	187	TWOLESSTILCSPLPSLSLVGYTKMKDLVFEWLEDAPAVOAEGTLTPOLFIIRDEKDLCG	246

[illegible]

DB 295 SLGVTLLTMTQASINSKLPVSYIKADVIGVCLAFIFGALLEYAVVNYGR--KE 352  
QY 363 FTRLRORORORLEEDIIQESR 384  
DB 353 FLR-KEKKKTRLDCCPSER 373

## RESULT 14

077295 PRELIMINARY; PRT; 453 AA.  
ID 077295;  
AC 077295;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE DROSGLUCL.  
GN GLUCL-ALPHA OR INDEFINITE OR CG7535.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Semenov E.P., Pak W.L.;  
RT "Identification of Drosophila Chloride Channel Gene by multiple  
RT posttranscriptional mRNA modifications";  
RL J. Neurochem. 0:0-0(0).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
DR EMBL: AA002232; CA003260.1; -;  
DR Flybase: FBgn0024963; GluCl-alpha.  
DR InterPro: IPR001188; GABA\_receptor.  
DR InterPro: IPR001175; GABA\_receptor.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_mem; 1.  
DR PRINTS: PR00252; NRIONCHANNEL.  
DR TIGRfams: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Transmembrane.  
SQ SEQUENCE 453 AA; 51905 MW; C0AD53FA6968C2A CRC64;

Query Match 33.0%; Score 742.5; DB 5; Length 453;  
Best local Similarity 43.1%; Pred. No. 2.4e-61;  
Matches 157; Conservative 58; Mismatches 120; Indels 29; Gaps 7;

QY 21 VLLRVALAKEEYKSGTSGSPSPSDFLKLNGRTSGYDARIIP--NFKGPPVNTCNI 77  
DB 9 ILTFASLCSASLANNAKVNFREREKKVLDQILG-AGKYDARIIPSGINGTDGPAVVRNI 67  
QY 78 FINSESVTKTMDYRVNVLKQONNDPRLSYREYDDSLDLPMSLDSIMKPDLPFANE 137  
DB 68 FVRSIKIDVTMEYSVQLTFREQWTDERLKFDDIGRLKYLTLTFANVWMPDLFFSNE 127  
QY 138 KGANFHEVTTDKLRIFFKNGVLSIRLTLLISCLMDLKNFPMIDQITQMOLESSIIC 197  
DB 128 KEGHFNIITIMPVYIRIFPNSGLYSIRISLTACPMNKLKPLDRQICSLM----- 180  
QY 198 SPLPSTLSISVGYTMKDLVEFMLEDAVAVOAGLTLPOFILDERDL--GCCGKHVNTGKF 256  
DB 181 -----ASYGWTNDLVFLMKEGDP-VGVVKNLHLPRFTL--EKLFDYCSKNTNGEY 230  
QY 257 TCIIEVGFHLEROMGYLLIOMYIPSLILVILSVSWFMINMDAAPARGLGITVLTMTQS 316  
DB 231 SCLKVDLFRRESYLLIYIPCCMLVIVSWVSFWLDOGAVPARSLGVTLLTMAIOT 290  
QY 317 SGRASLSPVSYVKAIDIMAVCLLFVFAALLEYAINVSR-----QKKEFTLRRLR 369  
DB 291 SGINASLSPVSYTKAIDVMTGCLTFVFGALLEFALVAVASRSGSKNANMKHESMKKKRR 350  
QY 370 QRRQ 373  
DB 351 DLEQ 354

## RESULT 15

094900 PRELIMINARY; PRT; 456 AA.  
ID 094900;  
AC 094900;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Glutamate-gated chloride channel.  
GN GLUCL-ALPHA OR DROSGLUCL OR CG7535.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-OREGON R; TISSUE=HEAD;  
RC MEDLINE=96355334; PubMed=8702744;  
RA Cullis D.F., Pares P.S., Liu K.K., Schaeffer J.M., Arena J.P.;  
RT "Identification of a Drosophila melanogaster glutamate-gated chloride  
RT channel sensitive to the antiparasitic agent avermectin";  
RL J. Biol. Chem. 271:20187-20191(1996).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.  
DR EMBL: U58776; AAC47266.1; -;  
DR Flybase: FBgn0024963; GluCl-alpha.  
DR InterPro: IPR001188; GABA\_receptor.  
DR InterPro: IPR001175; GABA\_receptor.  
DR Pfam: PF02931; Neur\_chan\_LBD; 1.  
DR Pfam: PF02932; Neur\_chan\_mem; 1.  
DR PRINTS: PR00252; NRIONCHANNEL.  
DR TIGRfams: TIGR00860; LIC; 1.  
DR PROSITE: PS00236; NEUROTR\_ION\_CHANNEL; 1.  
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Transmembrane.  
SQ SEQUENCE 456 AA; 52347 MW; 4008E5765756B51C CRC64;

Query Match 32.4%; Score 730.5; DB 5; Length 456;  
Best local Similarity 42.3%; Pred. No. 3.2e-60;  
Matches 154; Conservative 61; Mismatches 120; Indels 29; Gaps 7;

QY 21 VLLRVALAKEEYKSGTSGSPSPSDFLKLNGRTSGYDARIIP--NFKGPPVNTCNI 77  
DB 10 ILTFASLCSASLANNAKVNFREREKKVLDQILG-AGKYDARIIPSGINGTDGPAVVRNI 68  
QY 78 FINSESVTKTMDYRVNVLKQONNDPRLSYREYDDSLDLPMSLDSIMKPDLPFANE 137  
DB 69 FVRSIKIDVTMEYSVQLTFREQWTDERLKFDDIGRLKYLTLTFANVWMPDLFFSNE 128  
QY 138 KGANFHEVTTDKLRIFFKNGVLSIRLTLLISCLMDLKNFPMIDQITQMOLESSIIC 197  
DB 129 KEGHFNIITIMPVYIRIFPNSGLYSIRISLTACPMNKLKPLDRQICSLM----- 181  
QY 198 SPLPSTLSISVGYTMKDLVEFMLEDAVAVOAGLTLPOFILDERDL--GCCGKHVNTGKF 256  
DB 182 -----ASYGWTNDLVFLMKEGDP-VGVVKNLHLPRFTL--EKLFDYCSKNTNGEY 231  
QY 257 TCIIEVGFHLEROMGYLLIOMYIPSLILVILSVSWFMINMDAAPARGLGITVLTMTQS 316  
DB 232 SCLKVDLFRRESYLLIYIPCCMLVIVSWVSFWLDOGAVPARSLGVTLLTMAIOT 291  
QY 317 SGRASLSPVSYVKAIDIMAVCLLFVFAALLEYAINVSR-----QKKEFTLRRLR 369  
DB 292 SGINASLSPVSYTKAIDVMTGCLTFVFGALLEFALVAVASRSGSKNANMKHESMKKKRR 351  
QY 370 QRRQ 373  
DB 352 DLEQ 355

Search completed: June 25, 2003, 17:16:58  
Job time : 54.0671 secs



---

**THIS PAGE BLANK (USPTO)**

	Query Match	39.6%	Score 650.2;	DB 9;	Length 1857;
	Best Local Similarity	70.1%;	Pident No. 1,1e-169;		
	Matches 943;	Conservative 0;	Mismatches 348;	Indels 54;	Gaps 3
Oy	63	CCTCCTCAGGGTGGCCCTTTGGCAAAAGNGAAGTCAAATCTGNAACCAAGGGGTCCAGCC	122		
Db	450	CTTCAGGAGCGGCTTCTCGCAAAGCATGACTCCAGGTCTGGAANAACAACCTTCACAGAC	509		
Oy	123	CATGTCCCCCTTGATTCTCTAGACAAACTTATGGGGCAACATTCTGATATGATGCCAG	182		

[illegible]

**THIS PAGE BLANK (USPTO)**

1323 AGGAGAAACCGCGAACTCTACGTGACTGAGCCAGAGATTGACACCATCTCCG 1322  
1596 TGGAGATGCTATCAAGAGAAAGTGTGTGACCGGCAAAAGATTGACACATATCTCG 1655  
1323 GGCTCTCTTCCCTTCACTTCTTCTCAATCTCTCTGAGGTGTCTATTAAGT 1382  
1656 ACCTCCCTTCCCTTCCCTTCTTCTCAATCTTCTTCTGATCATCAAGAT 1715  
1383 GCTATGCTGAGAGATATCCACCAG 1407  
1716 CATTCGGCATGAGATGTCACAG 1740

RESULT 14  
AF268375 1744 bp mRNA linear MAN 22-MAY-2001  
LOCUS AF268375  
DEFINITION Bos taurus glycine receptor alpha 1 subunit (Glr1) mRNA, complete  
ACCESSION AF268375  
VERSION AF268375.1 GI:10180958  
KEYWORDS  
SOURCE Bos taurus.  
ORGANISM Bos taurus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
Bovidae; Bovinae; Bos.  
1 (bases 1 to 1744)  
Pierce, K.D., Handford, C.A., Morris, R., Vafa, B., Dennis, J.A.,  
Healy, P.J., and Schofield, P.R.  
A nonsense mutation in the alpha subunit of the inhibitory glycine  
receptor associated with bovine myoclonus  
Mol. Cell. Neurosci. 17 (2), 354-363 (2001)

JOURNAL MEDLINE  
PUBMED 11178872  
2 (bases 1 to 1744)  
Pierce, K.D., Handford, C.A., Morris, R., Vafa, B., Dennis, J.A.,  
Healy, P.J., and Schofield, P.R.  
Direct submission  
Submitted (17-MAY-2000) Neurobiology Program, Garvan Institute of  
Medical Research, 384 Victoria Street, Sydney, NSW 2010, Australia  
LOCATION/Qualifiers  
1.1744  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
1.1744  
/gene="Glr1"  
229.1602  
/note="Glr1s: alternatively spliced"  
/codon\_start=1  
/product="glycine receptor alpha 1 subunit"  
/protein\_id="AF268375.1"  
/db\_xref="GI:10180958"  
/translation="MSENTLRLLYMETIVFESLAASKAEARASRPMSPDELK  
LMGRSGYDARLPNFKPVPVSCNIFINSFSAIETMDRAVIFLQOONDRRLA  
YNEYDDSLDLPMDVOCIMOLFANKEGAPHEITDNTKLNLSRGNVLSIRI  
TTLACPMDLKDFPMDOVOCIMOLFANKEGAPHEITDNTKLNLSRGNVLSIRI  
BEKDLRYCTKHYNTGKFCIEARFHLEROMGYLLIOMTIPSLIYLSTISWIMDA  
APARVGLITTYLITWITSSGSRASLPKYSYKVIDIWAACLVLFVSLALEYAVNF  
VSRQKELFRFRKRRHRSKPMNLNFOEDAEGRFNSAYMGACACLOAKDGIKVG  
ANNSNTTNPAPSPKSPKEMRLFIORAKKIDKISRGPPMAFLIFNMYIYKIVR  
REDVHNO"

BASE COUNT 396 a 508 c 451 g 389 t  
ORIGIN

Query Match 39.4%; Score 646.8; DB 4; Length 1744;  
Best Local Similarity 71.2%; Pred. No. 9, 6e-169;  
Matches 956; Conservative 0; Mismatches 317; Indels 69; Gaps 5;

DB 89 AGGAGTCAAACTGGAACCAAGGCTCCAGCCATGCCCCCTGTGATTCTCAGACA 148  
299 AGGAGCTGAGCTGCTCGCTCTTCCAGCCATGACCCGTCCTCCGATTCCTGGATA 358

149 AACTATGAGGCGAATCTGATATGATGCCAGATTGCGCCAAATTTTAAAGGCCAC 208  
359 AACTATGAGGCGAATCTGATATGATGCCAGATTGCGCCAAATTTTAAAGGCCAC 418  
209 CCGTGAAGCTGACCTGCAACATCTTCATCAACAGTTGAGCTCGCTCAACCAAGCACAA 268  
419 CAGTGAATGTCAGCTCAACATTTTCATCAACAGCTGGTTCATTTGCTGAGAACACTA 478  
269 TGGACTACCGGCTGAATCTCTTCCGCGCAACAGTGAATGACCCAGCCCTGCTTAC 328  
479 TGGACTACCGGCTGAATCTCTTCCGCGCAACAGTGAATGACCCAGCCCTGCTTAC 538  
329 GAGATATCTGATGACTCTCTGACCTGATCCCTCAAGCTGATCTATCTGGAAGC 388  
539 ATGATATACCGGATGATCTCTGATCCCTCAAGCTGATCTATCTGGAAGC 598  
389 CAGACCTCTTCTTCTGCTAATGAGAAAGGGCCCACTTCCATGAGGTGACCAAGCACAA 448  
599 CTGACCTGCTTCTTCTGCTAATGAGAAAGGGCCCACTTCCATGAGGTGACCAAGCACAA 658  
449 ACTTACTGCGGCTCTTCAAGATGGGATGTCTGTACAGATCAAGCTGACCTCATTT 508  
659 AACTGCTGAGATCTCCCGGAAAGGGAATGTCTTACAGCATTTAGATCAACCTGACAT 718  
509 TGTCTGCGGCTGATGAGACCTGCAAGATCCCATGAGCATCCAGACGTCAGTACAGC 568  
719 TGCCCTGCGGCTGATGAGACCTGCAAGATCCCATGAGCATCCAGACGTCAGTACAGC 778  
569 TTGAGAGCTCATCATCTACTGACAGCCCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCT 628  
779 TGGAAAGCT-----TTGATATCA 796  
629 CCATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688  
797 CCATGAAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 853  
689 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748  
854 GACTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
749 ACTACAAACAGGAGAAATTCACCTGACAGTAAAGTTTACCTGAGAGGAGAGATG 808  
914 ATTACAAACAGGAGAAATTCACCTGACAGTAAAGTTTACCTGAGAGGAGAGATG 973  
809 GGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868  
974 GGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1033  
869 CTTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928  
1034 CTTTCTGATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1093  
929 TCACATGACCAACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988  
1094 TCACATGACCAACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1153  
988 AGGCAATGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1048  
1154 AGGCAATGACATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213  
1049 ATGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1108  
1214 AGCGCGGCTCAACTTGT 1273  
1109 AGAGCGCC-----AACGCTGAGAGAGATATCATCAAGAAAGTC 1150  
1274 GAAAGCATCAACAAGAGCCCATGTTGATTTCTCCAGAGATGAGGTGAGAGAGGCGC 1333  
1151 GTTTTATTTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207  
1334 GCTTCAACTTTCGCGCTTACGAGGATGAGGCGCCAGCTGCTTACGAGCCAAAGATGATCT 1393

**THIS PAGE BLANK (USPTO)**

```

/protein_id="CAC16687.1"
/db_xref="GI:11322384"
/translation="MTHTSLIINDCINSCRYHVFSGFESISPLCELYLMOKMSPS
DPLDKRMGRSGDARIKPKNGKPVNVCNFINSGSTETTMKRLNRLNVEFROOM
DPLATSEIPDASLDLDPMSLMDIKWPDLEFANEKANFEVTDNKLRIENGNVL
YSIRLRLILSCPMDLNFPMDIQCTMQLESFGYTMNDLLEFVLSNDPVOVADLTLP
OFVLEKEKIDGYCTKHNTGKTCIEVKEFLEHROMGYLLIOMIPSLIYLISWSEW
INDDAPARVGLITVLTMTTOSGSRASLPRYSYKALDIYMAVCLLIVFALLLEY
AANFVSRHKEFIRLRKROREIDEDIVRESGFGGLGLOCTODGAVES
SVAPPPVOVLYLDEAVKRRKRPDRAKRIDTISRANFPLIFLNFVMTTYVLRHE
DIRANP
1891. 2045
/gene="g1yr alpha"
BASE COUNT 524 a 501 c 460 g 560 t
ORIGIN
Alignment Scores:
Pred. No.: 8,72e-168 Length: 2045
Score: 1749.50 Matches: 334
Percent Similarity: 84.91% Conservative: 26
Best Local Similarity: 78.77% Mismatches: 25
Query Match: 77.69% Indels: 39
DB: Gaps: 4
US-10-075-846-4 (1-431) x DREA04970 (1-2045)
OY 9 LeuSerPheLeuLeuLeuLeuLeuProGlyLeuValLeuLeuValAlaLeuAla 28
DB 610 CTCTGCTTCTCTATCTCTG----- 630
OY 29 LysGluGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
DB 631 -----CAGAACCTATGTCCTCCATCGGACTTCTG 660
OY 49 AspLysLeuMetGlyArgThrSerGlyThrAspAlaGlyLeuArgProAspPheLeu 68
DB 661 GACAGCGATGAGGAGCAGAACCTCGCGATGATGCTCCATCAGACCCCACTTCAAGA 720
OY 69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88
DB 721 CCACACGTAACGTCACCTGTAACATCTTCAACAGTTTGGATCTATTACAGACACA 780
OY 89 ThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
DB 781 ACAATGATATACAGGCTTAACGCTTCTTACGGCAGCAGTGAACGACCTAGACTGCC 840
OY 109 TyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerIleTrp 128
DB 841 TACAGTGAATATCCCGATGATCTCTAGACTTGACCTTCTATGTTGACTCCATATGG 900
OY 129 LysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGluValThrThrAsp 148
DB 901 AAACCTGACTGTTTTCCTTATATAGAAAGTGGCACTTCCATGAAGTCCACCAACAC 960
OY 149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuThrSerIleArgLeuThrLeu 168
DB 961 AACAGAGTGAAGGATCTTTCAGAAATGGAATGCTTTCACAGCATCAGGCTACACATC 1020
OY 169 IleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMet 188
DB 1021 ATCCGTCTGTGCTCCATGACTTGAAGATTTCCCAATGAGACATTCAGACCTGTACCATG 1080
OY 189 GlnLeuGluSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208
DB 1081 CAGCTAGAA-----AGCTTTGGC 1098
OY 209 TyrThrMetLysAspLeuValPheGluTrpLeuGluAspAlaProAlaValGlnValAla 228
DB 1099 TACACCATGAACGATCTTTCGAGTGGCTTTCGATTAACCT---GTCCAGAGTTGGC 1155
OY 229 GluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuLysCysCysThr 248
DB 1156 GATGACTGACTTCTCTCAGTTTGTACTAAAGAGAGAAAGATCTCGGCTACTGCACT 1215

```

```

OY 249 LysHisTyrAsnThrGlyLysPheThrCysIleGluValLysPheHisLeuGluArgGln 268
DB 1216 AACCACTACAAACAGAGTAATATACCTGATGAGGTGAAGTTCACTAGAGACAG 1275
OY 269 MetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeuIleValIleLeuSerTrp 288
DB 1276 ATGGCGTACATCTCTCATTCAGATGATGATCCACAGCTGCTGATTCATTTGTCTGG 1335
OY 289 ValSerPheThrIleAsnMetAspAlaAlaProAlaArgValGlyLeuGlyIleThrThr 308
DB 1336 GTGCTTCTGATCAATATGATGATGCGCACCGGCTCGGTTGGGTATCACCACC 1395
OY 309 ValLeuThrMetThrThrGlnSerSerGlySerArgAlaSerLeuProLysValSerTyr 328
DB 1396 GTGCTGACATGACACACCGACAGCTGTTCAAGACCTCGGTACCAAGTGTCTTAC 1455
OY 329 ValLysAlaIleAspIleTrpMetAlaValCysLeuLeuPheValPheAlaIleLeu 348
DB 1456 GTGAAGCCCATTTGACATCTGATGCTGTTGTCTGCTGATGCTGCTGCTGCTGCTG 1515
OY 349 GluTyrAlaIleAlaIleAsnPheValSerArgGlnHisLysGluPheIleArgLeuArg 368
DB 1516 GAGTATCAGCAGATCACTGCTGCTCAAGCAGCAGCAGGATTCATCAGACTCCGCAAG 1575
OY 369 ArgGlnArgArgGlnArgLeuGluGluAspIleIleGlnGlnSerArg---PheTyrPhe 387
DB 1576 AAGCAGGAGGAGGAGGAGATAGAGAGACCTTGTGAGAGAAACCGTGCTTTACTTC 1635
OY 388 ArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaIleArgAspGlyGlyProMetGluGlySer 407
DB 1636 CGAGGTACGAGCAGCTGCGCATTTGCTTCAAAACAAAGAGGAGCAGAGTGGAGGGTCC 1695
OY 408 GlyIleTyrSerProGlnProAlaProLeuLeuArgGluGlyLeuThrThrArgLys 427
DB 1696 AGTATATTCGCCCAACCACTTCAGTCAAGTGTCTTACGACGAGAGCCGTCGGAGAG 1755
OY 428 LeuTyrValAsp 431
DB 1756 CGCTTCGTGAGC 1767
RESULT 5
LOCUS HSGLYRA1 1857 bp mRNA linear PRI 28-MAY-1993
DEFINITION "HSGLYRA1, alpha-2 strychnine binding subunit of inhibitory glycine receptor mRNA."
ACCESSION X52008.1 GI:31848
VERSION X52008.1
KEYWORDS glycine receptor; inhibitory glycine receptor; strychnine binding.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1857)
AUTHORS Greeninglugh,G., Schmieden,V., Schofield,P.R., Seeburg,P.H., Sidique,T., Mohandas,T.K., Becker,C.M. and Betz,H.
TITLE Alpha subunit variants of the human glycine receptor: primary structures, functional expression and chromosomal localization of the corresponding genes
JOURNAL EMBO J 9 (3), 771-776 (1990)
MEDLINE 90183975
PUBMED 2155780
FEATURES
source
Location/Qualifiers
1..1857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/translation="MTHTSLIINDCINSCRYHVFSGFESISPLCELYLMOKMSPS
DPLDKRMGRSGDARIKPKNGKPVNVCNFINSGSTETTMKRLNRLNVEFROOM
DPLATSEIPDASLDLDPMSLMDIKWPDLEFANEKANFEVTDNKLRIENGNVL
YSIRLRLILSCPMDLNFPMDIQCTMQLESFGYTMNDLLEFVLSNDPVOVADLTLP
OFVLEKEKIDGYCTKHNTGKTCIEVKEFLEHROMGYLLIOMIPSLIYLISWSEW
INDDAPARVGLITVLTMTTOSGSRASLPRYSYKALDIYMAVCLLIVFALLLEY
AANFVSRHKEFIRLRKROREIDEDIVRESGFGGLGLOCTODGAVES
SVAPPPVOVLYLDEAVKRRKRPDRAKRIDTISRANFPLIFLNFVMTTYVLRHE
DIRANP
1891. 2045
/gene="g1yr alpha"
BASE COUNT 524 a 501 c 460 g 560 t
ORIGIN
Alignment Scores:
Pred. No.: 8,72e-168 Length: 2045
Score: 1749.50 Matches: 334
Percent Similarity: 84.91% Conservative: 26
Best Local Similarity: 78.77% Mismatches: 25
Query Match: 77.69% Indels: 39
DB: Gaps: 4
US-10-075-846-4 (1-431) x DREA04970 (1-2045)
OY 9 LeuSerPheLeuLeuLeuLeuProGlyLeuValLeuLeuValAlaLeuAla 28
DB 610 CTCTGCTTCTCTATCTCTG----- 630
OY 29 LysGluGluValLysSerGlyThrLysGlySerGlnProMetSerProSerAspPheLeu 48
DB 631 -----CAGAACCTATGTCCTCCATCGGACTTCTG 660
OY 49 AspLysLeuMetGlyArgThrSerGlyThrAspAlaGlyLeuArgProAspPheLeu 68
DB 661 GACAGCGATGAGGAGCAGAACCTCGCGATGATGCTCCATCAGACCCCACTTCAAGA 720
OY 69 ProProValAsnValThrCysAsnIlePheIleAsnSerPheSerValThrLysThr 88
DB 721 CCACACGTAACGTCACCTGTAACATCTTCAACAGTTTGGATCTATTACAGACACA 780
OY 89 ThrMetAspTyrArgValAsnValPheLeuArgGlnGlnTrpAsnAspProArgLeuSer 108
DB 781 ACAATGATATACAGGCTTAACGCTTCTTACGGCAGCAGTGAACGACCTAGACTGCC 840
OY 109 TyrArgGluTyrProAspAspSerLeuAspLeuAspProSerMetLeuAspSerIleTrp 128
DB 841 TACAGTGAATATCCCGATGATCTCTAGACTTGACCTTCTATGTTGACTCCATATGG 900
OY 129 LysProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPheHisGluValThrThrAsp 148
DB 901 AAACCTGACTGTTTTCCTTATATAGAAAGTGGCACTTCCATGAAGTCCACCAACAC 960
OY 149 AsnLysLeuLeuArgIlePheLysAsnGlyAsnValLeuThrSerIleArgLeuThrLeu 168
DB 961 AACAGAGTGAAGGATCTTTCAGAAATGGAATGCTTTCACAGCATCAGGCTACACATC 1020
OY 169 IleLeuSerCysLeuMetAspLeuLysAsnPheProMetAspIleGlnThrCysThrMet 188
DB 1021 ATCCGTCTGTGCTCCATGACTTGAAGATTTCCCAATGAGACATTCAGACCTGTACCATG 1080
OY 189 GlnLeuGluSerSerIleLeuCysSerProLeuProSerLeuSerLeuSerValGly 208
DB 1081 CAGCTAGAA-----AGCTTTGGC 1098
OY 209 TyrThrMetLysAspLeuValPheGluTrpLeuGluAspAlaProAlaValGlnValAla 228
DB 1099 TACACCATGAACGATCTTTCGAGTGGCTTTCGATTAACCT---GTCCAGAGTTGGC 1155
OY 229 GluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLysAspLeuLysCysCysThr 248
DB 1156 GATGACTGACTTCTCTCAGTTTGTACTAAAGAGAGAAAGATCTCGGCTACTGCACT 1215

```

THIS PAGE BLANK (USPTO)



```

/codon_start=1
/product="Inhibitory glycine receptor"
/protein_id="CAA36257.1"
/db_xref="GI:31849"
/db_xref="SWISS-PROT:P23416"
/translation="MKROLVNLITLALFAFLEETNHRFAECKDHDSRSGKPSQTLSP
SDFKLKMGRTSGYDARLRPNKGPVNTCNIFINSFSEVETMDKRVNLFKROW
NDSRLAYSEYPDSDLDPSMDSIKPDLPEFANEKGFHDVTNDKLRISKNGV
LYSIRLTLELSCPMDLKNPMDVOTCTMOLSEFGYTMNDLIEFMLSDEGVYAEGLTL
POFLKEKEELGCTKHYNTKPTCTEVEFHLEROMGYKLIOMYIPSLIVLISVSE
WINMDAPARVALGITTYVTITMTOSGSAISPKSYVAIDIMAVCLLPFPAALLE
YAAVNFVRKHKEFLIRRRQRKNEEDVITRESRNFSEYGMGCHLOVKDSTAVKAT
PAMPPEPPQEDAIKKKFDRAKRIDTISRAPFLAFLENIIFYWITYKIRHEDVH
KK"
BASE COUNT      526 a      416 c      417 g      498 t
ORIGIN
Alignment Scores:
Pred. No.:      4,94e-161      Length:      1857
Score:          1682.50      Matches:      324
Percent Similarity: 86.10%      Conservative: 29
Best Local Similarity: 79.02%      Mismatches: 36
Query Match:      74.71%      Indels:      21
DB:              9      Gaps:      5
US-10-075-846-4 (1-431) x HSGLYRA1 (1-1857)
QY      24  ATGVALA1ALEUAlaLysGlu---GluValLysSerGlyThrLysGlySerGlnPromet 42
DB      454 AGGACGGCTTTCTTCGCAAGAACCATGACCTCCAGGTGGAACCAACCTTCACGACCCCTA 513
QY      43  SerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTyrAspAlaArgile 62
DB      514 TTCCTTCAGATTTCCTGGACAGTTAATGGGAAGACATCAGATATGATGCAAGAATC 573
QY      63  ArgProAsnPheLysGlyProproValAsnValThrCysAsnIlePheIleAsnSerPhe 82
DB      574 AGCCCAAAATTTTAAAGTCTCTCAGTAAAGCTTACTGCAATTTTATCAACAGCTTTT 633
QY      83  SerSerValThrLysThrMetLaspTyrArgValAsnValPheLeuArgGlnGlnTrp 102
DB      634 GGATCAGTCAAGAAACGACCATGACCTCCAGGTGGAATATTTTCTGACACACAGTGG 693
QY      103 AsnAspProArgLeuSerTyrArgGluTyrProAspAspSerLeuAspLeuAspProSer 122
DB      694 AATGATTACAGCGCTGCGTACGTAGTACCCAGATGATGATCCCTGCACTGGACCATCC 753
QY      123 MetLeuAspSerIleTrrPlyProAspLeuPhePheAlaAsnGluLysGlyAlaAsnPhe 142
DB      754 ATCTGAGACTCCATTGGAAACAGATTTGCTTTCCTGCAATAGAGAAGGCTGCCAATTC 813
QY      143 HisGluValThrThrAspAsnLysLeuLeuArgIlePhePheAsnGlnAsnValLeuTyr 162
DB      814 CAGCAGTCCACCACTGACAAACAAATTCATCGGATTTCCGAAATAGGCAAGTGGCTTAC 873
QY      163 SerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsnPhePrometLasp 182
DB      874 AGTATCAGACTACCTTACCTTATCCCTGATGAGTGAAGAATCTTCGATGAGAT 933
QY      183 IleGlnThrCysThrMetGlnLeuGluSerSerIleLeuLysSerProLeuProSer 202
DB      934 GTCCAGACCTGATCAATGACGCTGAG----- 960
QY      203 LeuSerLeuSerValGlyTyrThrMetLysAspLeuValPheGlnTrpLeuGluAspAla 222
DB      961 -----AGTTTGGGTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1011
QY      223 ProAlaValGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArgAspGluLys 242
DB      1012 CCA---GTGCAAGTGTGCTGAAGATTGACCTGCCAGATTATTTTGAAGAAGAAGAG 1068
QY      243 AspLeuGlyCysCysThrLysThrLysThrAsnThrGlyLysPheThrCysIleGluValLys 262
DB      -----

```

```

DB      1069 GAACCTGGCTACTGTACAAAGCAGTACACACTGGAAAGTTTACCTGCATTGAGCTCAAG 1128
QY      263 PheH1sLeuGluArgGlnMetGlyTyrTyrLeuIleGlnMetTyrIleProSerLeuLeu 282
DB      1129 TTTCACCTGGAGCGCAATATGGATATTTATGATCCAGATATACATCCCAAGCTGCTT 1188
QY      283 ILeValIleLeuSerTyrValSerPheTrrPileAsnMetAspAlaAlaProAlaArgVal 302
DB      1189 AATGAAATTTTGTCTCTGGGTTCTTTTGATTAATATGATGACAGCCCTCCAGGGTGC 1248
QY      303 GlyLeuGlyIleThrThyValLeuThrMetThrTrgInSerSerGlySerArgAlaSer 322
DB      1249 GCACCGGGCATCCACAGATTTAAGATGACACCCAGAGTTGAGCTCCAGGGCATCT 1308
QY      322 LeuProLysValSerTyrValLysAlaIleAspIleTrrPrometAlaValCysLeuLeuPhe 342
DB      1309 CTGCCAAAGGCTCTCTATGTAAAGCGATTGACATCTGATGCGGTGCTTCTGTTT 1368
QY      343 ValPheAlaAlaLeuLeuGlnLysValAlaAlaIleAsnPheValSerArgGlnHisLysGlu 362
DB      1369 GTGTTTGCTGCTTACTGGAATACGAGCGGGAACCTTGCTCCAGGCAACAGAGAG 1428
QY      363 PheIleArgLeuArgArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArg 382
DB      1429 TTCCTGGCGCTCCGAGAGAGACAGAGAGCAAGTAAAGAGAGACCTTACTGCTGAA 1488
QY      383 SerArgPheTyrPheArgGlyTyrGlyLeuGlnLysCysLeuGlnAlaArgAspGlyGly 402
DB      1489 AGTCGTTTAATTTATAGCGTTATGAGATGGGTGCTGCTCTCAAGTGAAGAATGAGACA 1548
QY      403 PrometGlnGlySerGlyIleTrrSer---ProGlnProProAlaProLeuLeuArgGlu 421
DB      1549 GGTGTCAAGGCCACACCTCCCAACCATCTCCCAACAGCCCA-----AAAGAT 1596
QY      422 GlyGluThrThrArgLysLeuTyrValAsp 431
DB      1597 GGAGATGCTATCAAGAAAGTTGTGGAC 1626

```

RESULT 6  
RNO310837  
LOCUS Rattus norvegicus mRNA for glycine receptor alpha 2 precursor.  
DEFINITION Rattus norvegicus mRNA for glycine receptor alpha 2 precursor.  
ACCESSION AJ310837  
VERSION AJ310837.1 GI:13548660  
KEYWORDS glycine receptor alpha 2 precursor.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 1359)  
AUTHORS Beato M., Groot-Kormelink P.J., Colquhoun D. and Slyvlotz L.G.  
TITLE Concentration dependence of single channel currents through rat recombinant alpha 1 glycine receptors  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1359)  
AUTHORS Groot-Kormelink P.J.  
TITLE Direct Substitution  
JOURNAL Submitted (30-MAR-2001) Groot-Kormelink P.J., Department of Pharmacology, The School of Pharmacy, 29/39, Brunswick Square, London, WC1N 1AX, UNITED KINGDOM  
FEATURES  
source  
1. 1359  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley males"  
/db\_xref="taxon:10116"  
/tissue\_type="whole brain"  
/clone\_lib="rat brain marathon-ready cDNA"  
/dev\_stage="8-12 weeks"  
1. 1359  
/function="ion channel"  
/codon\_start=1  
/evidence="experimental"

**THIS PAGE BLANK (USPTO)**

XX	WT1: 2000-656166/63.	
DR	P-PSDB: AAB19336.	
XX	Novel nucleic acid sequence encoding human ataxia protein for screening	
PT	compounds useful for treating disorders relating to mutations in ataxia	
PT	gene	
XX	Claim 2; Page 18-20; 47pp; English.	
PS		
XX	The present sequence encodes a human ataxia protein. The ataxia	
CC	protein and polynucleotides are useful for diagnosing and treating	
CC	disorders related to ataxia. Ataxia gene sequences are useful in	
CC	gene therapy, and as diagnostic tools or reagents for identifying and	
CC	characterizing genetic defect involved in the disorders and diseases	
CC	related to ataxia.	
XX		
SO	Sequence 1600 BP: 378 A; 440 C; 379 G; 403 T; 0 other:	
	Query Match 91.7%; Score 1503.8; DB 21; Length 1600;	
	Best Local Similarity 96.6%; Pred. No. 0;	
	Matches 1585; Conservative 0; Mismatches 12; Indels 44; Gaps 3	
QY	1 ATGACAACTCTGTTGCTTCCTGCAACCTCTCCCTTCCTTCCTTCCTGGAACCTCGCAGGCGAG 60	
DB	1 ATGACAACTCTGTTGCTTCCTGCAACCTCTCCCTTCCTTCCTTCCTGGAACCTCGCAGGCGAG 60	
QY	61 GTCTCTCCACAGGAGTGCCCTTG6CAAAAGAGAAAGTCAATCTGGAACCAAGGGGTCCAG 120	
DB	61 GTCTCTCCACAGGAGTGCCCTTG6CAAAAGAGAAAGTCAATCTGGAACCAAGGGGTCCAG 120	
QY	121 CCCATGTCCTCCCTCTGATTTCTCTAGACAACTTATG6GGCGCAACTCTGATATGATGCC 180	
DB	121 CCCATGTCCTCCCTCTGATTTCTCTAGACAACTTATG6GGCGCAACTCTGATATGATGCC 180	
QY	181 AGGATTCGGCCCAATTTTAAAGGCCACCCTCGGAACGTGACTGCAACATCTTCATCAAC 240	
DB	181 AGGATTCGGCCCAATTTTAAAGGCCACCCTCGGAACGTGACTGCAACATCTTCATCAAC 240	
QY	241 AGTTTCACGCTCCGTCACCAAGACACCAATGACACTACGGGTGAATGTCTCTTGGCGCAA 300	
DB	241 AGTTTCACGCTCCATATCAACAGACCAATGACACTACGGGTGAATGTCTCTTGGCGCAA 300	
QY	301 CAGTGAATGAGCCACGCGCTGCTCTACCGAGAAATATCTGATGACTCTCTGGAACCTCGAT 360	
DB	301 CAGTGAATGAGCCACGCGCTGCTCTACCGAGAAATATCTGATGACTCTCTGGAACCTCGAT 360	
QY	361 CCCTTCATGCTGAGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTPATAGAAAGGGGCC 420	
DB	361 CCCTTCATGCTGAGACTCTATCTGGAAGCCAGACCTCTTCTTGTCTPATAGAAAGGGGCC 420	
QY	421 AACCTTCACAGGAGTGACCAAGGACAAACAAATGACGCGATCTTCAAGAAATGGGAATG 480	
DB	421 AACCTTCACAGGAGTGACCAAGGACAAACAAATGACGCGATCTTCAAGAAATGGGAATG 480	
QY	481 CTGTACAGCATCAGGCTGACCTCACTTTTGTCTCCTGATGAGACTCAAGAACTTCCCG 540	
DB	481 CTGTACAGCATCAGGCTGACCTCACTTTTGTCTCCTGATGAGACTCAAGAACTTCCCG 540	
QY	541 ATGACATTCACAGCTGACGATGACGACTTGAGAGCTCATCACTCTGACGCCCTCTG 600	
DB	541 ATGACATTCACAGCTGACGATGACGACTTGAGAGCTCATCACTCTGACGCCCTCTG 600	
QY	601 CCATCTCTGTCACTTTCAAGTGGCTACACCAACCAAGAAACCTGCTGTATGAGCGTGGAA 660	
DB	601 CCATCTCTGTCACTTTCAAGTGGCTACACCAACCAAGAAACCTGCTGTATGAGCGTGGAA 660	
QY	661 GATGCTCTGCTGTCTCAAGTGGCTAGAGGCTGACTCTGCCAGTTTATCTTGGCGGAT 720	
DB	619 GATGCTCTGCTGTCTCAAGTGGCTAGAGGCTGACTCTGCCAGTTTATCTTGGCGGAT 720	
QY	721 GAGAAGATCTAAGGCTGTTGTACCAAGCACTACAACAGGAAATTCACCTGATCGAG 780	
DB	721 GAGAAGATCTAAGGCTGTTGTACCAAGCACTACAACAGGAAATTCACCTGATCGAG 780	

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (usps)**

us-10-075-846-4.p2n.rng

Accession	Gene	Protein	Accession	Gene	Protein
Dp	GAGAAGATCTAGAGCTGTTGTACCAAGACATCAACAACAGAGGAATAATTACCTGCATCGAG	1023	Dp	GAGAAGATCTAGAGCTGTTGTACCAAGACATCAACAACAGAGGAATAATTACCTGCATCGAG	1023
Qy	261 VallysPhehlSLeuGluArgGlnMeGlyTyrTyrLeuIleGlnMetTyrIleProSer	280	Qy	261 VallysPhehlSLeuGluArgGlnMeGlyTyrTyrLeuIleGlnMetTyrIleProSer	280
Dp	1024 GTAAGATTTCACCTGGAAAGCGCAGATGGGCTACATCATCGATTGATCATGATCCCCAGC	1083	Dp	1024 GTAAGATTTCACCTGGAAAGCGCAGATGGGCTACATCATCGATTGATCATGATCCCCAGC	1083
Qy	281 LeuLeuIleValIleLeuSerTyrValSerPheThrPheLeuMetAspAlaIleProIle	300	Qy	281 LeuLeuIleValIleLeuSerTyrValSerPheThrPheLeuMetAspAlaIleProIle	300
Dp	1084 CTACTCATCGATCCTCTGCTGGGTCTCTTGATGATCAACATGATGATGCTGCTGCC	1143	Dp	1084 CTACTCATCGATCCTCTGCTGGGTCTCTTGATGATCAACATGATGATGCTGCTGCC	1143
Qy	301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArg	320	Qy	301 ArgValGlyLeuGlyIleThrThrValLeuThrMetThrThrGlnSerSerGlySerArg	320
Dp	1144 CGTGTGGGCGTGGGCGATCACCACCGGTGCTCAACATGACACACACAGCTGTGGCTCCGG	1203	Dp	1144 CGTGTGGGCGTGGGCGATCACCACCGGTGCTCAACATGACACACACAGCTGTGGCTCCGG	1203
Qy	321 AlaSerLeuProLysValSerTyrValLysAlaIleAspIleTyrMetAlaValCysLeu	340	Qy	321 AlaSerLeuProLysValSerTyrValLysAlaIleAspIleTyrMetAlaValCysLeu	340
Dp	1204 GCCTCTTGGCTTAAGCTGCTCTACCTGAAGGCATATCGACATCTGGATGGCTGTGCTGG	1263	Dp	1204 GCCTCTTGGCTTAAGCTGCTCTACCTGAAGGCATATCGACATCTGGATGGCTGTGCTGG	1263
Qy	341 LeuPheValPheAlaIleLeuLeuGlnLysThrAlaAlaIleAsnPheValSerArgGlnHis	360	Qy	341 LeuPheValPheAlaIleLeuLeuGlnLysThrAlaAlaIleAsnPheValSerArgGlnHis	360
Dp	1264 CTCTTTGTTGGTGCGCTTGCTGGAGATGCTGCATTAATTTGTTCTGCTGAGAT	1323	Dp	1264 CTCTTTGTTGGTGCGCTTGCTGGAGATGCTGCATTAATTTGTTCTGCTGAGAT	1323
Qy	361 LysGluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluLysPheIle	380	Qy	361 LysGluPheIleArgLeuArgArgArgGlnArgArgGlnArgLeuGluLysPheIle	380
Dp	1324 AAAGATTCATACGACGCTTCGAAAGAGCGACAGAGCCGCCCAACGCTTGAGAGATATCATC	1383	Dp	1324 AAAGATTCATACGACGCTTCGAAAGAGCGACAGAGCCGCCCAACGCTTGAGAGATATCATC	1383
Qy	381 GlnGluSerArgPheThrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaIleArgAsp	400	Qy	381 GlnGluSerArgPheThrPheArgGlyTyrGlyLeuGlyHisCysLeuGlnAlaIleArgAsp	400
Dp	1384 CAAAGAAATCGCTTATTTCCTGCTGATGGCTTGGGCCCACTGCCTCAGCAGAGAAT	1443	Dp	1384 CAAAGAAATCGCTTATTTCCTGCTGATGGCTTGGGCCCACTGCCTCAGCAGAGAAT	1443
Qy	401 GlyGlyProMetGluGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg	420	Qy	401 GlyGlyProMetGluGlySerGlyIleTyrSerProGlnProProAlaProLeuLeuArg	420
Dp	1444 GGAGGTCCAAATGGAAGTTCTGGCATTTATTAAGTCCCAACCTCCAGCCCTCTTCTAAAG	1503	Dp	1444 GGAGGTCCAAATGGAAGTTCTGGCATTTATTAAGTCCCAACCTCCAGCCCTCTTCTAAAG	1503
Qy	421 GlnGlyGluThrThrArgLysLeuTyrValAsp	431	Qy	421 GlnGlyGluThrThrArgLysLeuTyrValAsp	431
Dp	1504 GAAGAGAAACACCGCGGAACCTACGTGAGC	1536	Dp	1504 GAAGAGAAACACCGCGGAACCTACGTGAGC	1536
RESULT 2					
AAC61678					
Accession	Gene	Protein	Accession	Gene	Protein
Dp	AAC61678	standard; DNA; 1600 BP.	Dp	AAC61678	standard; DNA; 1600 BP.
AC	AAC61678;		AC	AAC61678;	
XX			XX		
XX	19-FEB-2001 (first entry)		XX	19-FEB-2001 (first entry)	
DE	CDNA sequence encoding a human ataxia protein.		DE	CDNA sequence encoding a human ataxia protein.	
XX			XX		
XX	Human; ataxia; gene therapy; ss.		XX	Human; ataxia; gene therapy; ss.	
OS	Homo sapiens.		OS	Homo sapiens.	
XX			XX		
XX	Key	Location/Qualifiers	XX	Key	Location/Qualifiers
FT	CDS	1..1254	FT	CDS	1..1254
FT		/tag="a	FT		/tag="a
FT		/product="ataxia protein"	FT		/product="ataxia protein"
XX	WO200058461-A1.		XX	WO200058461-A1.	
PN			PN		
PD	05-OCT-2000.		PD	05-OCT-2000.	
XX			XX		
PF	23-MAR-2000; 2000WO-EP02600.		PF	23-MAR-2000; 2000WO-EP02600.	
XX			XX		
PR	26-MAR-1999; 99EP-0106343.		PR	26-MAR-1999; 99EP-0106343.	
XX			XX		
PA	(RAPF/) RAPFOLD-HOERBRAND G.		PA	(RAPF/) RAPFOLD-HOERBRAND G.	
XX			XX		
PI	Rapfold-Hoerbrand G;		PI	Rapfold-Hoerbrand G;	
XX			XX		
XX	WPI; 2000-656166/63.		XX	WPI; 2000-656166/63.	
DR	P-PSDB; AAB19336.		DR	P-PSDB; AAB19336.	
XX			XX		

[illegible]

**THIS PAGE BLANK (USPTO)**



```

Db      679 GAGAGAGATCTAGGCTGTTGTTACCAACACACAGGGAATTCACCTGATCGAG 738
Qy      261 VALLVSPHEHISLEUGLUAARGIMETGLYTYRILEUILEGINMETYRILEPROSER 280
Db      739 GTAAGATTTCACCTGAGACGGGAGATGGGCTACTATCTGATTGATGATACATCCCCACG 798
Qy      281 LeuLeuIleValIleLeuSerTrpValSerPheTrpIleAsnMetAspAlaIleProAla 300
Db      799 CTACTCATGTCATCCTGCTCTGCTGCTCTCTCTCTGATCAACATGATGCTGCCCTGCC 858
Qy      301 ArgValGlyLeuGlyIleThrTrpValIleuThrMetThrThcInserSerGlySerArg 320
Db      859 CGTGTGGCTGGGCAATCACACACCTGCTCACCATGACACACAGAGCTCTGGCTCCGG 918
Qy      321 AlaSerLeuProValSerTrpValIleValIleAspIleTrpMetAlaValCysLeu 340
Db      919 GCCTCTTTCCTAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 978
Qy      341 LeuPheValIleAlaAlaLeuLeuGluTrpAlaAlaIleAsnPheValSerArgIleHis 360
Db      979 CTCTTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1038
Qy      361 LysGluPheIleArgIleuArgArgArgGlnArgGlnArgGlnArgGlnArgGlnArg 380
Db      1039 AAGAAATTCATACGACTTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098
Qy      381 GlnGluSerArgPheTrpPheArgGlyTrpGlyLeuGlyHisCysLeuGlnAlaArgAsp 400
Db      1099 CAGAGAAAGTCGTTTCATTTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
Qy      401 GlyGlyPrometGluGlySerGlyIleTrpSerProGlnProAlaProLeuLeuArg 420
Db      1159 GGGAGTCATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1218
Qy      421 GlnGlyGluThrTrpArgLysLeuTrpValAsp 431
Db      1219 GAGGAGAAACACGCGGAGAACTCATGAGTGCAC 1251

RESULT 3
AAQ25722
ID      AAQ25722 standard; DNA; 657 BP.
XX
AC      AAQ25722;
XX
XX      08-DEC-1992 (first entry)
XX
XX      GABA-A receptor beta-subunit.
XX
XX      Expression plasmid; gamma-aminobutyric acid; OmpF; glycine; ss.
XX
XX      Synthetic.
XX
XX      JP04144683-A.
XX
XX      19-MAY-1992.
XX
XX      05-OCT-1990; 90JP-0267743.
XX
XX      05-OCT-1990; 90JP-0267743.
XX
XX      (MITU) MITSUBISHI KASEI CORP.
XX
XX      WPI; 1992-214122/26.
XX
XX      Prep. of N-terminal extracellular site protein - by culturing
XX      E.coli transformed by a plasmid comprising the tac promoter, ribosome
XX      binding site, etc.
XX
XX      Claim 1; Page 2; 12pp; Japanese.
XX
XX      The sequence given is the gamma-aminobutyric acid (GABA) A receptor
XX      beta-subunit. This sequence is used in an expression plasmid operably

```

```

CC      linked to the tac promoter, a ribosome binding sequence, the E. coli
CC      outer membrane protein OmpF signal peptide coding sequence and a
CC      sequence coding for the N-terminal extracellular site of either the
CC      gamma-aminobutyric acid (GABA) A receptor alpha-subunit or the glycine
CC      receptor alpha-subunit. This expression plasmid can be used to
CC      transform E. coli to produce an N-terminal extracellular site protein
CC      of ion channel direct binding type receptor.
XX
SQ      Sequence 657 BP; 171 A; 185 C; 156 G; 145 T; 0 other;

Alignment Scores:
Pred. No.:      1,86e-104      Length:      657
Score:          984.50         Matches:     182
Percent Similarity: 87.34%      Conservative: 18
Best Local Similarity: 79.48%      Mismatches:  14
Query Match:    43.72%         Indels:      15
DB:             13             Gaps:        2

US-10-075-846-4 (1-431) x AAQ25722 (1-657)
Qy      40 GlnPrometSerProSerAspPheLeuAspLysLeuMetGlyArgThrSerGlyTrpAsp 59
Db      16 AAGCCTATGTCACCCCTCGGACTTCCTCGATATGGAAGGACTTCTGGTATGAT 75
Qy      60 AlaArgIleArgProAsnPheLysGlyProproValAsnValThrCysAsnIlePheIle 79
Db      76 GCCAGATCAGACCCCACTTAAGGTCTCCTGTAACGTAGTGAACATCTTCATC 135
Qy      80 AsnSerPheSerSerValIleThrTrpThrMetLaspTrpArgValAsnAlaPheLeuArg 99
Db      136 AACAGCTTGTGTTCTATGCCGAGACACCATGAGTACAGGTCAACATCTTCTGAGG 195
Qy      100 GlnGlnTrpAsnAspProArgLeuSerTrpArgGlyTrpProAspAspSerLeuAspLeu 119
Db      196 CAGCACTGGAAGACGCCCGCTCGCTCCATGATATCCCGACGACTCTGACCTT 255
Qy      120 AspProSerMetLeuAspSerIleTrpLysProAspLeuPheAlaAsnGlySgly 139
Db      256 GACCCATTCATGTTGATTCATCTGGAAGCCCTGACTTCTTCTGCAATGAGAAGGGG 315
Qy      140 AlaAsnPheHisGluValThrTrpAspAsnLysLeuLeuArgIlePheLysAsnGlyAsn 159
Db      316 GCCCACTTCACGAGATCCACGAGACGACGACGAGCTGCTGAGATCTCCCGAAGCGCAC 375
Qy      160 ValLeuTrpSerIleArgLeuThrLeuIleLeuSerCysLeuMetAspLeuLysAsn 179
Db      376 GTCCTCTACAGATCAGATACACCTGACTCGGCTGCCCATGAGACTGGAAGATTTTC 435
Qy      180 PrometAspIleGlnThrCysThrMetGlnLeuGluSerSerIleLeuCysSerPro 199
Db      436 CCGATGAGCTACAGACATGATCATCACTGGA----- 471
Qy      200 LeuProSerLeuSerLeuSerValGlyTrpThrMetLysAspLeuValPheGluTrpLeu 219
Db      472 -----AGCTTGGTATACCATGACACCTCATCTTGAAGTGG--- 510
Qy      220 GluAspAlaProAlaValAlaGlnValAlaGluGlyLeuThrLeuProGlnPheIleLeuArg 239
Db      511 GAGAGACAGAGAGCTGTGCGAGTGGGAGATGAGCTGACCTGCTGCTGCTGCTGCTGCTG 570
Qy      240 AspGluLysAspLeuGlyCysThrLysHisTrpAsnThrGlyLysPheThrCysIle 259
Db      571 GAGAGAGAGATCTGAGTACTGACACCAAGCATCAACACAGGTAAATTCACCTGCAATT 630
Qy      260 GluValLysPheHisLeuGluArgGln 268
Db      631 GAGGCCGATTCACCTGGAAGCGCAG 657

RESULT 4
ID      ABI99254 standard; cDNA; 2404 BP.
XX
XX      ABI99254
XX
XX      ABI99254;

```

**THIS PAGE BLANK (USPTO)**